Revised December 1991

DYNAMIC LINEAR MODELS WITH LEADING INDICATORS

by

Jingxian Chen
The George Washington University

GWU/IRRA/Serial TR-91/4
September, 1991

Research Supported by

Contract N00014-85-K-0202
Office of Naval Research

and

Grant DAAL03-87-K-0056
U.S. Army Research Office

Grant AFOSR-89-0381
The Air Force Office of Scientific Research

and

The Ford Motor Company

DYNAMIC LINEAR MODELS WITH LEADING INDICATORS

by

Jingxian Chen

Abstract

This thesis proposes a dynamic linear model (DLM) to deal with the problem of forecasting with leading indicators. We call this type of a DLM as a dynamic linear model with leading indicators. Our approach expands the conventional one-dimension DLMs to the two-dimension case. Analyses of some real data sets which initially motivated us to explore our approach, are used as applications. For reasons of confidentiality they have been coded as Data Set One, Data Set Two and Data Set Three, respectively. Our approach has a much wider field of application, for instances, the two-dimension filter problems in image processing, and estimation problems related to Markov random fields.

1. INTRODUCTION AND OVERVIEW

The problem addressed in this thesis is to expand upon the dynamic linear model (DLM), such that it can deal with leading indicators. It is motivated by the analysis of three time series data sets listed in Tables 1.1 - 1.3, whose plots are shown in Figures 1.1 - 1.3. Each data set involves several series; for example, Data Set One consists of six series. From those plots, particularly from Figure 1.1, one can perceive that there may exist a pattern common to all the series within the set. For instance, all the six series of Data Set One reveal an 'S' shaped pattern. This phenomenon motivates us to think of introducing a leading indicator series into the DLM to improve the forecasts of a series of interest. For example, if we want to forecast series 6 of data set one, we may want to use data from not only series 6 but also series 1 to 5 to improve the forecasts of series 6.

Table 1.1 Data Set One

t	Series 1: Y_1 $y_{1,t}$	Series 2: Y_2 $y_{2,t}$	Series 3: Y_3 $y_{3,t}$	Series 4: Y_4 $y_{4,t}$	Series 5: Y_5 $y_{5,t}$	Series 6: Y_6 $y_{6,t}$
0	0	0	0	0	0	0
1	0.2	0.3	3.9	0.8	0.8	0.5
2	3.3	1.4	7.2	3.0	2.7	2.0
3	8.5	3.5	12.1	7.1	5.8	5.4
4	14.5	6.1	18.1	12.7	9.8	10.0
5	22.7	10.2	26.2	20.1	17.2	17.1
6	34.5	14.7	36.6	29.0	26.4	27.9
7	48.2	20.9	49.9	40.9	39.1	39.7
8	64.2	28.5	64.7	56.6	53.2	57.0
9	81.8	37.2	81.6	72.4	70.0	75.6
10	101.2	46.2	98.6	89.7	88.4	95.0
11	123.7	56.6	123.2	111.9	108.8	116.8
12	148.0	68.3	146.6	134.4	131.8	141.6
13	169.4	81.5	171.3	156.4	157.4	170.6
14	194.3	92.8	194.0	181.2	184.4	199.4
15	217.2	103.3	214.5	199.9	211.5	222.3
16	233.9	112.0	232.1	215.3	231.0	225.8
17	252.1	120.7	248.0	233.2	250.8	260.8
18	266.0	129.0	260.8	248.5	270.2	276.1
19	278.9	135.9	273.5	262.6	285.1	293.6
20	289.6	141.6	284.7	273.4	299.8	305.2
21	298.9	146.5	293.7	281.4	312.0	314.4
22	306.0	152.6	300.9	288.4	320.7	322.8
23	312.0	153.6	307.7	294.2	327.9	330.2
24	316.7	156.0	298.0	298.5	334.4	337.6
25	320.1	157.6	317.6	301.8	338.9	341.7
26	322.6	158.7	319.9	303.3	342.1	345.8
27	324.3	159.4	320.9	304.1	341.9	346.5
28	325.3	159.9	321.8	304.8	346.5	348.4
29	326.3	160.4	322.7	305.4	344.9	350.2
30	327.7	160.5	323.3	305.9	346.3	351.7
31	328.2	160.7	323.9	306.2	349.3	353.3
32	328.4	160.8	324.1	306.3	350.6	354.8
33	328.6	160.9	324.2	306.5	351.0	356.4
34	328.8	161.0	324.4	306.8	352.1	
35	328.9	161.1	324.5	307.0	352.7	
36	329.0	161.2	324.6	307.0	353.3	
37	329.1	161.2	324.7	307.1	353.3	
38	329.3	161.3	324.7	307.2	353.7	
39	329.0	161.3	324.7	307.3	352.2	

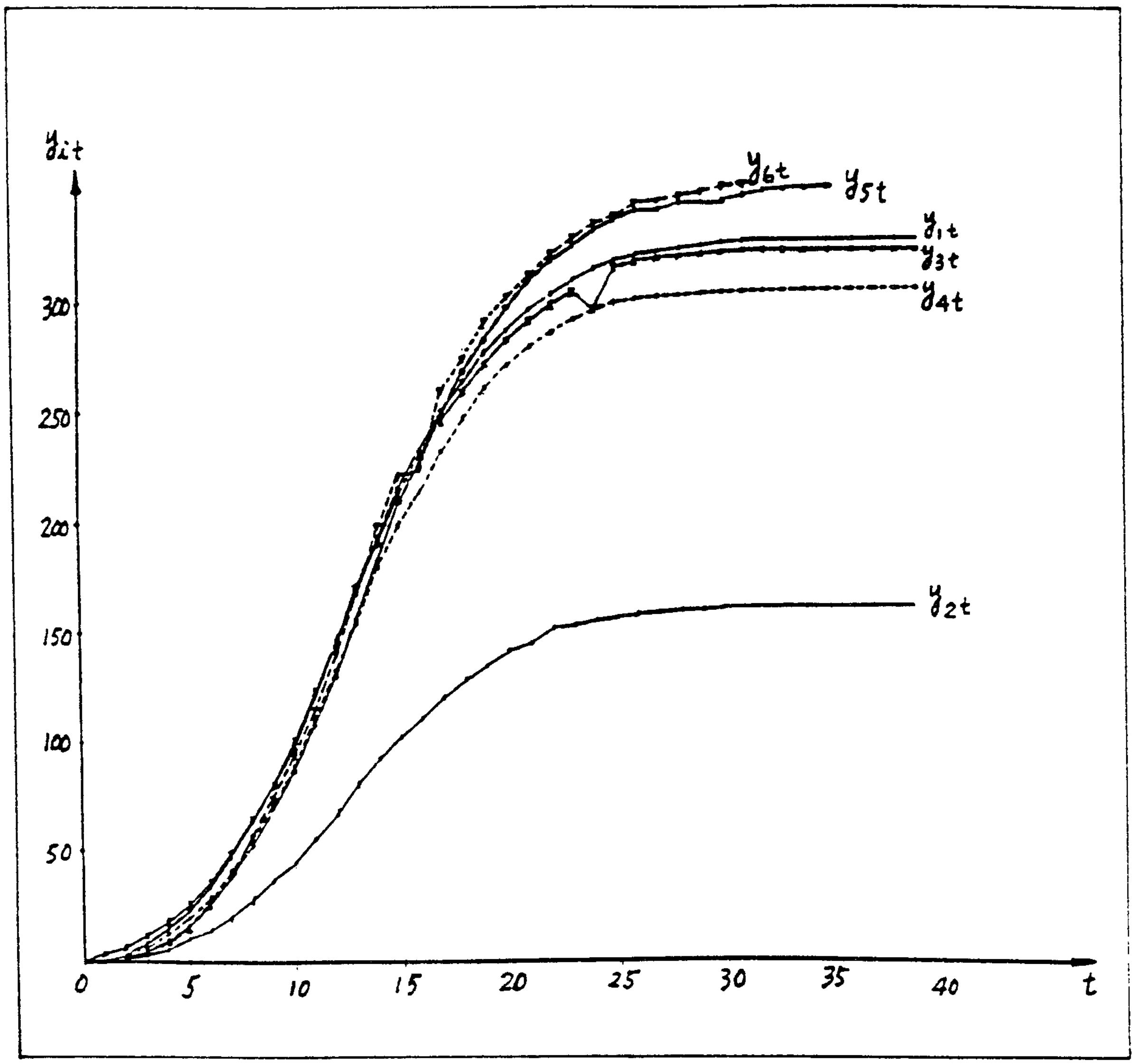

Figure 1.1 A Plot of Data Set 1

Table 1.2 Data Set Two

t	Series 1: Y_1 $y_{1,t}$	Series 2: Y_2 $y_{2,t}$	Series 3: Y_3 $y_{3,t}$
0	9.337	8.43	9.181
1	20.53	22.35	20.78
2	30.35	38.48	29.05
3	40.29	53.58	37.29
4	50.47	67.14	45.53
5	61.08	79.02	52.25
6	71.49	89.23	58.09
7	81.74	98.49	63.5
8	92.39	107.2	68.35
9	103.1	115.2	72.95
10	114	122.9	77.34
11	124.7	130	81.7
12	137.9	137.9	86.5
13	145.3	141.9	89.01
14	153.1	145.8	91.87
15	161.3	149.3	94.27
16	170.3	152.8	96.71
17	179.6	156.5	99.45
18	189	160.3	101.8
19	198.7	164.2	104.2
20	208.7	168.4	
21	218.7	172.8	
22	228.5	177.5	
23	238.8	182.8	
24	249.1	188.2	
25	258.8	193	
26	268.2	198.1	
27	278	203.4	
28	288.3	207.9	
29	299.1	212.8	
30	309.9	215.1	
31	321.5		
32	334		
33	346.6		
34	359.8		
35	372.2		
36	385.3		
37	395.9		
38	406.1		
39	416.1		

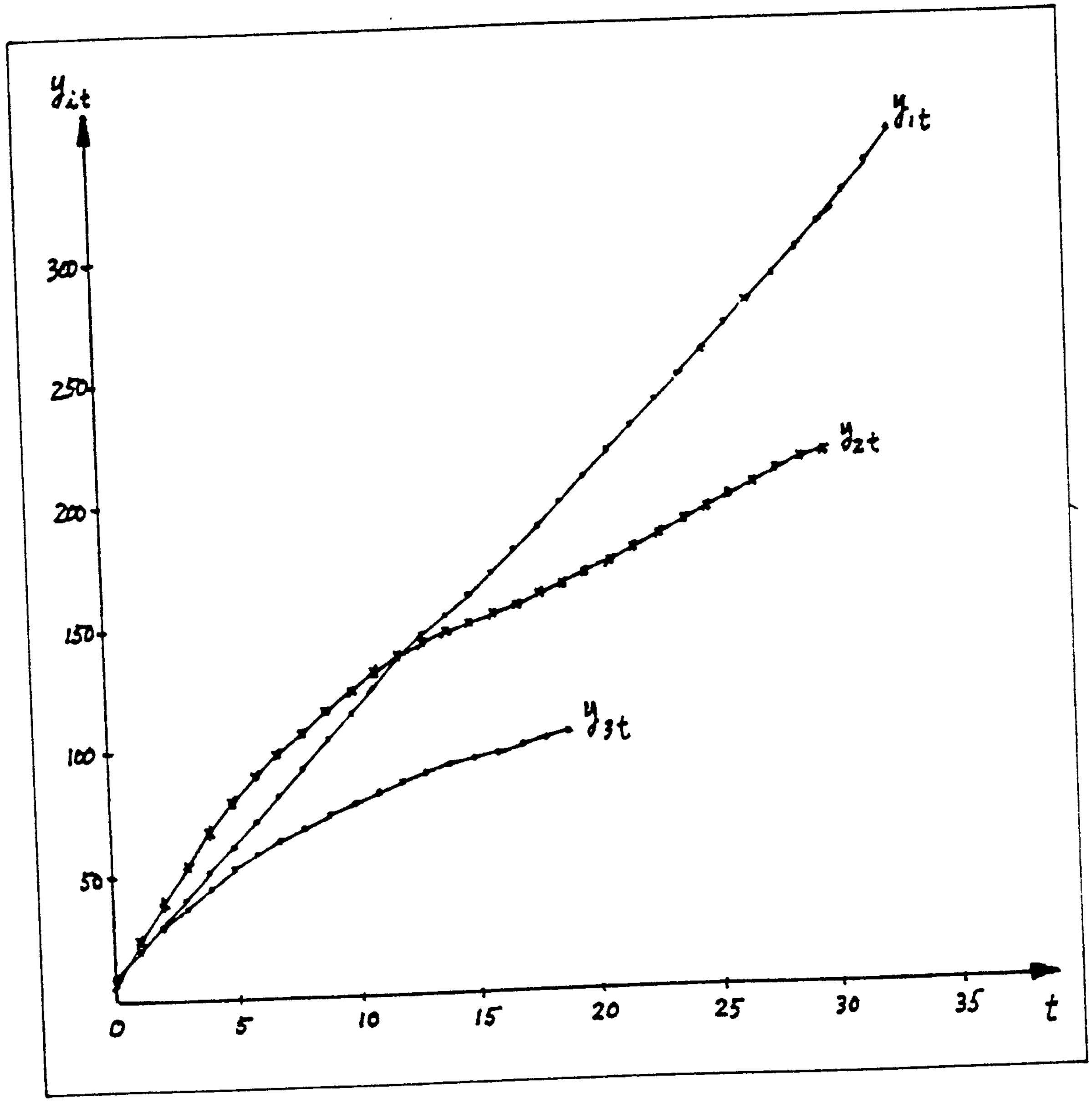

Figure 1.2 A Plot of Data Set 2

Table 1.3 Data Set Three

t	Series 1: Y_1 $y_{1,t}$	Series 2: Y_2 $y_{2,t}$	Series 3: Y_3 $y_{3,t}$
0	.862	.119	.131
1	3.467	.21	.266
2	6.36	.291	.355
3	8.902	.347	.45
4	10.93	.434	.543
5	12.46	.525	.63
6	13.78	.586	.693
7	14.96	.682	.752
8	16.05	.738	.814
9	17.15	.802	.845
10	18.48	.87	.915
11	19.8	.937	.974
12	21.21	.994	1.098
13	22.28	1.042	1.162
14	23.32	1.087	1.196
15	24.17	1.142	1.196
16	25.06	1.186	
17	25.76	1.244	
18	26.57	1.285	
19	27.39	1.339	
20	28.27	1.383	
21	29.09	1.403	
22	29.92	1.428	
23	30.74	1.55	
24	31.68	1.59	
25	32.53	1.662	
26	33.31	1.725	
27	33.97	1.725	
28	34.68		
29	35.42		
30	36.13		
31	36.98		
32	37.81		
33	38.76		
34	39.55		
35	40.32		
36	41.25		
37	42.08		
38	42.7		
39	42.92		

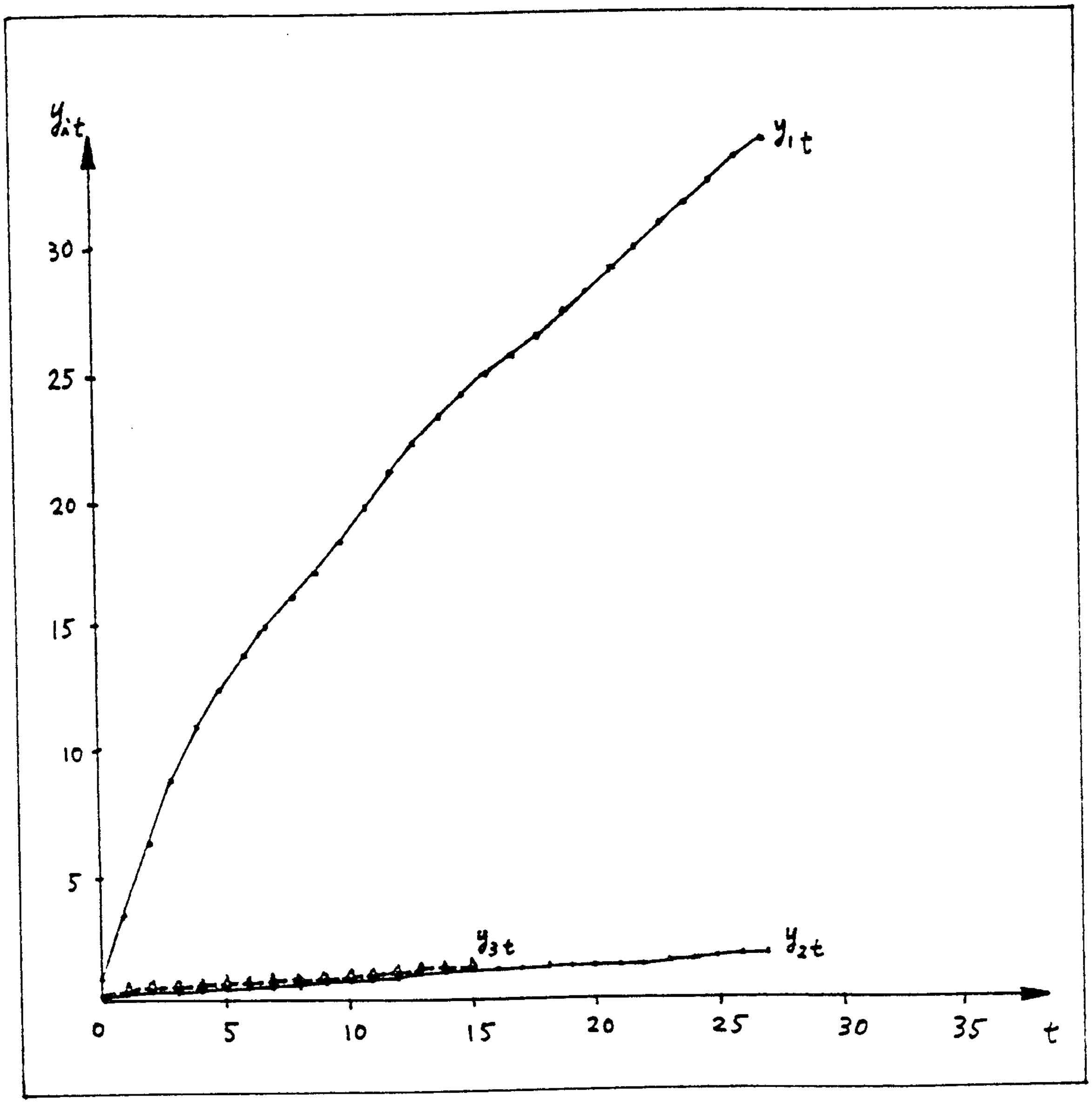

Figure 1.3 A Plot of Data Set 3

In Section 2 we overview the DLM. We propose the DLM with leading indicators in Section 3. Section 4 pertains to applications. In Section 5 we briefly discuss our future work. Details of the forecasting formulae of the DLM with leading indicators are in the Appendix.

2. OVERVIEW OF DLM

DLMs, also known as Kalman filter (KF) models, are models from which a large number of useful forecasting schemes can be derived as special cases [cf. Harrison and Stevens (1971, 1976) or West and Harrison (1989)]. In what follows, first we review the general model and forecasting formulae of the DLM, then we specify some particular KF models pertaining to processing our data analysis.

2.1 General Form and Forecasting Formulae of DLM

The DLM is often specified as

$$\begin{cases} Y_t = \mathbf{F}_t' \boldsymbol{\theta}_t + u_t , \\ \boldsymbol{\theta}_t = G_t \boldsymbol{\theta}_{t-1} + \mathbf{w}_t, \end{cases} \tag{2.1}$$

where

- $\boldsymbol{\theta}_t$ is the parameter vector of the system at time t and it is unobservable
- $\mathbf{F}_t$ is the known dynamic regression vector at time t,
- G_t is the known state evolution matrix at time t,
- u_t is the observation error at time t,
- $\mathbf{w}_t$ is the system error vector at time t,
- $Y_1, \ldots, Y_t \ldots$ is the observed series.

The first equation is called *observation equation* and the second one is called *system equation.*

Kalman (1960) has obtained the recursive equations and the predictive distributions for the DLM of form (2.1) under the Gaussian set-up. Kalman's results are summarized below.

Let "$X \sim N(\mu, \sigma^2)$" denote the fact that X has a Gaussian distribution with mean μ and variance σ^2.

If $u_t \sim N(0, \mathbf{U}_t)$, $\mathbf{w}_t \sim N(\mathbf{0}, \mathbf{W}_t)$, the sequences $\{u_t\}$ and $\{\mathbf{w}_t\}$ are mutually independent, the prior of $\boldsymbol{\theta}_0$ (before observing any data from the time series) is $N(\mathbf{m}_0, \mathbf{C}_0)$, then

i) $$(\boldsymbol{\theta}_t \mid \underset{\sim}{y}_t) \sim N(\mathbf{m}_t, \mathbf{C}_t), \tag{2.2}$$

where $\underset{\sim}{y}_t$ is the observation set $(y_1, y_2, \ldots, y_t)$; i.e. the realization of $Y_1, \ldots, Y_t$,

$$\mathbf{m}_t = \mathbf{G}_t \mathbf{m}_{t-1} + \mathbf{R}_t \mathbf{F}_t Q_t^{-1} (y_t - \mathbf{F}_t' \mathbf{G}_t \mathbf{m}_{t-1}),$$

$$\mathbf{C}_t = \mathbf{R}_t - \mathbf{R}_t \mathbf{F}_t \mathbf{F}_t' \mathbf{R}_t' Q_t^{-1},$$

$$\mathbf{R}_t = \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t' + \mathbf{W}_t,$$

$$Q_t = \mathbf{F}_t' \mathbf{R}_t \mathbf{F}_t + \mathbf{U}_t;$$

ii) $$(Y_{t+k} \mid \underset{\sim}{y}_t) \sim N(f_{t+k}, \sigma^2_{t+k}), \qquad \text{for } k=1, 2, \ldots \tag{2.3}$$

where $f_{t+k} = \mathbf{F}_{t+k}' \mathbf{G}_{t+k} \mathbf{G}_{t+k-1} \cdots \mathbf{G}_{t+1} \mathbf{m}_t$, and

$$\sigma^2_{t+k} = \mathbf{F}_{t+k}' \mathbf{G}_{t+k} \mathbf{G}_{t+k-1} \cdots \mathbf{G}_{t+1} \mathbf{C}_t \mathbf{G}_{t+1}' \cdots \mathbf{G}_{t+k-1}' \mathbf{G}_{t+k}' \mathbf{F}_{t+k}$$

$$+ \mathbf{F}_{t+k}' \left(\sum_{i=1}^{k-1} \left(\mathbf{G}_{t+k} \cdots \mathbf{G}_{t+i+1}\right) \mathbf{W}_{t+i} \left(\mathbf{G}_{t+k} \cdots \mathbf{G}_{t+i+1}\right)' + \mathbf{W}_{t+k} \right) \mathbf{F}_{t+k}$$

$$+ \mathbf{U}_{t+k}.$$

Smith and West (1983) have studied the case wherein $\mathbf{U}_t$ and $\mathbf{W}_t$ are not fully specified, and their results are given below.

If $(u_t \mid U) \sim N(0, U)$, and the prior of U is an inverted gamma with shape parameter $n_0/2$ and scale parameter $d_0/2$, $(\mathbf{w}_t \mid U) \sim N(\mathbf{0}, U\mathbf{W}_t^*)$, $(\boldsymbol{\theta}_0 \mid U) \sim N(\mathbf{m}_0, U\mathbf{C}_0^*)$, with $\mathbf{W}_t^*$ and

C_0^* specified, and if given U the sequences $\{u_t\}$ and $\{\mathbf{w}_t\}$ are assumed mutually independent, then

$$(\boldsymbol{\theta}_t \mid \underline{y}_t) \sim T_{n_t}(\mathbf{m}_t,\ \mathbf{C}_t)\ , \tag{2.4}$$

where $T_{n_t}(\mathbf{m}_t,\ \mathbf{C}_t)$ denotes the multivariate student-t distribution with n_t degrees of freedom, mode $\mathbf{m}_t$ and scale matrix $\mathbf{C}_t$, with

$$\mathbf{m}_t = \mathbf{G}_t\, \mathbf{m}_{t-1} + \mathbf{R}_t\, \mathbf{F}_t\, Q_t^{-1}\, (y_t - \mathbf{F}_t'\, \mathbf{G}_t\, \mathbf{m}_{t-1}),$$

$$\mathbf{C}_t = \frac{d_t/n_t}{d_{t-1}/n_{t-1}}\, (\mathbf{R}_t - \mathbf{R}_t\, \mathbf{F}_t\, \mathbf{F}_t'\, \mathbf{R}_t'/Q_t),$$

$$\mathbf{R}_t = \frac{d_{t-1}}{n_{t-1}}\, (\mathbf{G}_t\, \mathbf{C}_{t-1}^*\, \mathbf{G}_t' + \mathbf{W}_t^*),$$

$$Q_t = \frac{d_{t-1}}{n_{t-1}} + \mathbf{F}_t'\, \mathbf{R}_t\, \mathbf{F}_t\ ,$$

$$n_t = n_{t-1} + 1,$$

$$d_t = d_{t-1} + \frac{d_{t-1}}{n_{t-1}}\, (y_t - \mathbf{F}_t'\, \mathbf{G}_t\, \mathbf{m}_{t-1})^2/Q_t\ ,$$

$$\mathbf{C}_t^* = \frac{n_t}{d_t}\, \mathbf{C}_t.$$

In this case, the k-step ahead predictive distribution is also a univariate Student-t distribution with n_t degrees of freedom, mode f_{t+k} and scale σ_{t+k}^2 where

$$f_{t+k} = \mathbf{F}_{t+k}'\, (\mathbf{G}_{t+k}\, \mathbf{G}_{t+k-1} \cdots \mathbf{G}_{t+1})\, \mathbf{m}_t\ ,$$

$$\sigma_{t+k}^2 = \mathbf{F}_{t+k}'\, \mathbf{G}_{t+k}\, \mathbf{G}_{t+k+1} \cdots \mathbf{G}_{t+1})\, \mathbf{C}_t\, \mathbf{G}_{t+k}' \cdots \mathbf{G}_{t+k-1}'\, \mathbf{G}_{t+k}')\, \mathbf{F}_{t+k}$$

$$+ \frac{d_t}{n_t}\, (1 + \mathbf{F}_{t+k}'\, (\mathbf{W}_{t+k}^* + \sum_{i=1}^{k-1} (\mathbf{G}_{t+k} \cdots \mathbf{G}_{t+i+1})\, \mathbf{W}_{t+i}^* (\mathbf{G}_{t+i+1}' \cdots \mathbf{G}_{t+k}'))\, \mathbf{F}_{t+k}),$$

$$\text{for } k=1,\, 2,\, \ldots\ . \tag{2.5}$$

Hence, a general application of DLM to a practical forecasting problem consists of two tasks: i) a specification of a particular form of (2.1), and ii) a specification of the distributions of the errors and the prior of θ_0 to set the forecasting conditions in one of the preceding two schemes.

2.2 Growth Models

Because of the trend feature in all the three data sets, we will concentrate on the specification of growth models.

Growth models are a class of special DLMs. A growth model can be described by two parameters θ_t and β_t at time t. The former is often called "level" and the latter called "change of level" or "slope." In terms of model (2.1), the growth model is defined as

$$\boldsymbol{\theta}_t = \begin{bmatrix} \theta_t \\ \beta_t \end{bmatrix}, \ \mathbf{F}_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{w}_t = \begin{bmatrix} v_t \\ w_t \end{bmatrix} \ \text{and} \ \mathbf{G}_t = \begin{bmatrix} 1 & 1 \\ 0 & k_t \end{bmatrix};$$

Alternatively, it can be written as

$$\begin{cases} Y_t = \theta_t + u_t, \\ \theta_t = \theta_{t-1} + \beta_{t-1} + v_t, \\ \beta_t = k_t \beta_{t-1} + w_t . \end{cases} \tag{2.6}$$

Specifically, we use the following three growth models to analyze our data sets.

1) *Linear growth model*

   This is the simplest case of growth models. Here, K_t is always specified as 1. This model has been found suitable for dealing with data having a linear upward trend; for instance y_{2t} in data set 3.

2) *Segmented linear growth model*

Examine the plots of data set 2 (Figure 1.2). All the three series show an upward trend, but the trend changes at its slope about t=13. Therefore, K_t is specified as

$$K_t = \begin{cases} 1, & \text{if } t \neq 13, \\ 0.9, & \text{if } t = 13. \end{cases}$$

In general, if the change point of the trend is at t_0, the specification of K_t is of the form

$$K_t = \begin{cases} 1, & \text{if } t \neq t_0, \\ \text{Positive constant } (\neq 1), & \text{at } t = t_0. \end{cases}$$

3) *S-shaped growth model*

This model is suitable for data set 1; such data shows a trend in three stages: at first, for $t \leq 12$, the growth rate appears increasing, then it appears to be constant, and finally, for $t \geq 16$, it decrease and tends to zero. In effect, the trend is an S-shaped one. Thus for data set 1, K_t is specified as follows:

$$K_t = \begin{cases} 1.15, & \text{if } t \leq 12, \\ 1, & \text{if } 13 \leq t \leq 15, \\ 0.85, & \text{if } t \geq 16. \end{cases}$$

The specification of K_t for a general 'S' - shaped trend is

$$K_t = \begin{cases} C_1 \ (\text{constant} > 1), & \text{if } t \leq t_1, \\ 1 \quad , & \text{if } t_1 \leq t \leq t_2, \\ C_2 \ (\text{constant} \in (0, 1)), & \text{if } t > t_2, \end{cases}$$

where t_1, t_2 are terminals of stages 1 and 2.

3. DLM WITH LEADING INDICATORS

Now we start our expansion of the DLM. Before doing so, we need to introduce a more elaborate notation. Let

$Y_{\ell,t}$ denote an observation on the ℓ-th series at time t,

$\theta_{\ell,t}$ is a state parameter, standing for the level of the ℓ-th series at time t,

$\beta_{\ell,t}$ is the other state parameter, reflecting the change of the level of the ℓ-th series at time t,

$u_{\ell,t}$ is the observation error of the ℓ-th series at time t,

$v_{\ell,t}$ and $w_{\ell,t}$ are the system errors of the ℓ-th series at time t, corresponding to parameter $\theta_{\ell,t}$ and $\beta_{\ell,t}$, respectively.

3.1 The Forms of DLM with Leading Indicators

Without loss of generality, let $\{Y_{2,t}\}$ be the series of interest and $\{Y_{1,t}\}$ be the leading indicator series. Assuming a growth model, we introduce a weight γ to incorporate the effect of the indicator series in the series of interest. The following two forms can be considered.

1) Weighted on level, $\theta_{\cdot,t}$

For the leading indicator series, we have, as before

$$\begin{cases} Y_{1,t} = \theta_{1,t} + u_{1,t}, \\ \theta_{1,t} = \theta_{1,t-1} + \beta_{1,t-1} + v_{1,t}, \\ \beta_{1,t} = K_t\beta_{1,t-1} + w_{1,t}; \end{cases} \tag{3.1a}$$

and for the series of interest, we have

$$\begin{cases} Y_{2,t} = \theta_{2,t} + u_{2,t}, \\ \theta_{2,t} = \gamma\theta_{2,t-1} + (1-\gamma)\,\theta_{1,t} + v_{2,t}. \end{cases} \tag{3.1b}$$

2) Weighted on slope $\beta_{\cdot,t}$

$$\begin{cases} Y_{1,t} = \theta_{1,t} + u_{1,t}, \\ \theta_{1,t} = \theta_{1,t-1} + \beta_{1,t-1} + v_{1,t}, \\ \beta_{1,t} = K_t\,\beta_{1,t-1} + w_{1,t}; \end{cases} \tag{3.2a}$$

$$
\begin{cases}
Y_{2,t} = \theta_{2,t} + u_{2,t}, \\
\theta_{2,t} = \theta_{2,t-1} + \beta_{2,t-1} + v_{2,t}, \\
\beta_{2,t} = \gamma K_t \, \beta_{2,t-1} + (1-\gamma)\, \beta_{1,t} + w_{2,t}.
\end{cases}
\tag{3.2b}
$$

As said before, for Data Sets One, Two and Three, the linear growth model, the segmented linear growth model and the S-shaped growth model are adopted respectively. For all three data sets, the second form performs better than the first one.

The assignment of weight γ is subjective; it is based on the forecaster's belief in the chosen model to describe the series of interest, and the potential benefit from the use of information from the indicator series. For example, if $\gamma=1$, Y_1 does not play any role and (3.1a) is reduced to a steady model, whereas (3.2a) is reduced to a conventional growth model.

In general, if we partition the parameter vector $\boldsymbol{\theta}_{.,t}$ as $\boldsymbol{\theta}_{.,t} = \begin{bmatrix} \boldsymbol{\theta}^{(1)}_{.,t} \\ \boldsymbol{\theta}^{(2)}_{.,t} \end{bmatrix}$ with $\boldsymbol{\theta}^{(1)}_{.,t}$ being the state parameters specific to the series of interest and $\boldsymbol{\theta}^{(2)}_{.,t}$ being those state parameters containing "pattern information" common for all series. Then the general expression of the DLM is of the form

$$
\begin{aligned}
Y_{\ell,t} &= \mathbf{F}'_t \, \boldsymbol{\theta}_{\ell,t} + u_{\ell,t}, \\
\boldsymbol{\theta}_{\ell,t} &= \mathbf{H}_{\ell,t} \, \boldsymbol{\theta}_{\ell,t} + \mathbf{J}_{\ell,t} \, \boldsymbol{\theta}_{\ell-1,t} + \mathbf{w}_{\ell,t},
\end{aligned}
\tag{3.3}
$$

where

$$
\mathbf{H}_{\ell,t} = \begin{bmatrix} \mathbf{G}^{(11)}_t & \mathbf{G}^{(12)}_t \\ \mathbf{G}^{(21)}_t & \boldsymbol{\Gamma}_\ell \mathbf{G}^{(22)}_t \end{bmatrix}, \quad \mathbf{J}_{\ell,t} = \begin{bmatrix} 0 & 0 \\ 0 & (\mathbf{I}-\boldsymbol{\Gamma}_\ell) \end{bmatrix},
$$

$$\mathbf{G}_t = \begin{bmatrix} \mathbf{G}_t^{(11)} & \mathbf{G}_t^{(12)} \\ \mathbf{G}_t^{(21)} & \mathbf{G}_t^{(22)} \end{bmatrix}$$ partitioned corresponding to

$$\boldsymbol{\theta}_{\cdot,t} = \begin{bmatrix} \boldsymbol{\theta}_{\cdot,t}^{(1)} \\ \boldsymbol{\theta}_{\cdot,t}^{(2)} \end{bmatrix},$$

$\boldsymbol{\Gamma}_\ell$ is a diagonal matrix with entries taking values between 0 and 1, when ℓ=1, $\boldsymbol{\Gamma}_1 = \mathbf{I}$.

3.1.1 Relationship to a 2-Dimensional KF Model

A two-dimension KF model is often defined as

$$\begin{aligned} Y_{\ell,t} &= \mathbf{F}_t' \boldsymbol{\theta}_{\ell,t} + u_{\ell,t}, \\ \boldsymbol{\theta}_{\ell,t} &= \mathbf{H}_{\ell,t} \boldsymbol{\theta}_{\ell,t-1} + \mathbf{J}_{\ell,t} \boldsymbol{\theta}_{\ell-1,t} + \mathbf{K}_{\ell,t} \boldsymbol{\theta}_{\ell-1,t-1} + \mathbf{w}_{\ell,t} \end{aligned} \tag{3.4}$$

[cf Habibi (1972), Woods and Radewan (1977), Woods and Ingle (1981), Katayama and Kosaka (1979), etc.]. In this sense (3.3) can be regarded as a special case of (3.4) with $\mathbf{K}_{\ell,t} \equiv \mathbf{0}$. From now on, we call this expansion form of the DLM as the *DLM with leading indicators.*

3.2 Inference and Forecasting Formulae

Let us consider a simple example of the DLM with leading indicators to demonstrate our method by which we solve the inference and forecasting problem of the DLM with leading indicators. Then, we write down the general inference and forecasting formulae.

3.2.1 An Example of the DLM with Leading Indicators

Suppose that observations $y_{1,1}$, $y_{1,2}$, $y_{2,1}$, $y_{2,2}$ are available and predictions $Y_{2,3}$ and $Y_{2,4}$ are requested. Assume the suitable model is (3.2) and the priors for $\theta_{1,0}$ and $\beta_{1,0}$ are $(\theta_{1,0}\,|\,U) \sim N(m_{1,0}, UC_{1,0})$, $(\beta_{1,0}\,|\,U) \sim N(b_{1,0}, U\sigma_{1,0})$, $(u.,t\,|\,U) \sim N(0, U)$, $(v., t\,|\,U) \sim N(0, UV)$, $(w.,t\,|\,U) \sim N(0, UW)$, the prior of U is an inverted gamma with parameters $\frac{n_0}{2}$ and $\frac{d_0}{2}$. We also assume the same independence conditions as we deal with the DLM [cf. West and Harrison (1989)]. Besides, we assume that $\theta_{2,0} = \theta_{1,0} + s_1$, $\theta_{2,0} = \beta_{1,0} + s_2$ where $s_1$and s_2 are known constants. Under these assumptions follow from (3.2a)

$$\theta_{1,0} = \theta_{1,0}\,, \qquad\qquad \beta_{1,0} = \beta_{1,0}\,,$$

$$\theta_{1,1} = \theta_{1,0} + \beta_{1,0} + v_{1,1}\,, \qquad\qquad \beta_{1,1} = k_1\beta_{1,0} + w_{1,1}\,,$$

$$\theta_{1,2} = \theta_{1,0} + (1+k_1)\beta_{1,0} + v_{1,1} + w_{1,1} + v_{1,2}\,,$$

$$\beta_{1,2} = k_2k_1\beta_{1,0} + k_2w_{1,1} + w_{1,2}\,,$$

....

$$\theta_{1,4} = \theta_{1,0} + (1+k_1(1+k_2(1+k_3)))\beta_{1,0} + v_{1,1} + (1+k_2(1+k_3))w_{1,1} + v_{1,2}$$
$$+ (1+k_3)w_{1,2} + v_{1,3} + w_{1,3} + v_{1,4}\,,$$

$$\beta_{1,4} = k_4\cdots k_1\beta_{1,0} + k_4k_3k_2w_{1,1} + k_4k_3w_{1,2} + k_4w_{1,3} + w_{1,4}\,,$$

$$Y_{1,1} = \theta_{1,0} + \beta_{1,0} + v_{1,1} + u_{1,1}\,,$$

$$Y_{1,2} = \theta_{1,0} + (1+k_1)\beta_{1,0} + v_{1,1} + w_{1,1} + v_{1,2}\,.$$

Alternatively, it can be represented in the matrix form $(\boldsymbol{\Theta}_1', \mathrm{Y}_{1,1}, \mathrm{Y}_{1,2})' = \mathbf{L}\,\mathbf{Z}$,

where $\quad \boldsymbol{\Theta}_1 = (\theta_{1,0}\ \beta_{1,0}\ \theta_{1,1}\ \beta_{1,1}\ \theta_{1,2}\ \beta_{1,2}\ \theta_{1,3}\ \beta_{1,3}\ \theta_{1,4}\ \beta_{1,4})'$,

$$\mathbf{L} = \begin{bmatrix}
1 & & & & & & & & & & & \\
0 & 1 & & & & & & & & & & \\
1 & 1 & 1 & & & & & & & & & \\
0 & k_1 & 0 & 1 & & & & & & & & \\
1 & 1+k_1 & 1 & 1 & 1 & & & & & & & \\
0 & k_2k_1 & 0 & k_2 s & 0 & 1 & & & & & & \\
1 & 1+k_1+k_1k_2 & 1 & 1+k_2 & 1 & 1 & 1 & & & & & \\
0 & k_3k_2k_1 s & 0 & k_3k_2 & 0 & k_3 & 0 & 1 & & & & \\
1 & 1+k_1(1+k_2(1+k_3)) & 1 & 1+k_2+k_2k_3 & 1 & 1+k_3 & 1 & 1 & 1 & & & \\
0 & k_4k_3k_2k_1 & 0 & k_4k_3k_2 & 0 & k_4k_3 & 0 & k_4 & 0 & 1 & & \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \\
1 & 1+k_1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix},$$

$$\mathbf{Z} = [\theta_{1,0}\ \beta_{1,0}\ v_{1,1}\ w_{1,1}\ v_{1,2}\ w_{1,2}\ v_{1,3}\ w_{1,3}\ v_{1,4}\ w_{1,4}\ u_{1,1}\ u_{1,2}\,]'\,.$$

Assume that $(\mathbf{Z} \mid \mathrm{U}) \sim \mathrm{N}(\boldsymbol{\mu}, \mathrm{U}\Sigma_0)$, $\mathrm{U} \sim \mathrm{IG}(\frac{n_0}{2}, \frac{d_0}{2})$. Then, by the properties of the multivariate normal distribution, we can obtain

$$(\theta_{1,0}, \beta_{1,0}, \cdots, \theta_{1,4}, \beta_{1,4} \mid \mathrm{U}, y_{1,1}, y_{1,2}) \sim \mathrm{N}(\mathbf{M}_1, \mathrm{U}\Sigma_1),$$

$$(\theta_{1,0}, \beta_{1,0}, \cdots, \theta_{1,4}, \beta_{1,4} \mid y_{1,1}, y_{1,2}) \sim \mathrm{T}_{n_0+2}(\mathbf{M}_1, \frac{d^{(1)}}{n_0+2}\,\Sigma_1),$$

$$(\mathrm{U} \mid y_{1,1}, y_{1,2}) \sim \mathrm{IG}(\frac{n_0+2}{2}, \frac{d^{(1)}}{2}),$$

where $\mathbf{M}_1$, Σ_1, $d^{(1)}$ can be calculated out (for brevity, we omit the formulae for calculating $\mathbf{M}_1$, Σ_1, $d^{(1)}$).

Similarly, (3.2b) offers a tool to obtain the relationship between ($\theta_{2,0}$, $\beta_{2,0}$, $\cdots$, $\theta_{2,4}$, $\beta_{2,4}$, $Y_{2,1}$, $Y_{2,2}$, $Y_{2,3}$, $Y_{2,4}$) and ($\theta_{1,0}$, $\beta_{1,0}$, $\cdots$, $\theta_{1,4}$, $\beta_{1,4}$, $v_{2,1}$, $w_{2,1}$, $\cdots$, $v_{2,4}$, $w_{2,4}$, $u_{2,1}$, $\cdots$, $u_{2,4}$), and we can get the predictive distribution $(Y_{2,3}, Y_{2,4} \mid y_{1,1}, y_{1,2}, y_{2,1}, y_{2,2})$ in which the leading indicator series $\{Y_{1,t}\}$ is incorporated.

3.2.2 General Formulae

Let n be the forecast horizon;

let $\underset{\sim}{Y}_{\ell}^{(1)} = [Y_{\ell,1}, \cdots , Y_{\ell,t}]'$ and $\underset{\sim}{y}_{\ell}^{(1)} = \underset{\sim}{y}_{\ell,t} = [y_{\ell,1}, \cdots , y_{\ell,t}]'$ be a realization of $\underset{\sim}{Y}_{\ell}^{(1)}$,

$\underset{\sim}{Y}_{\ell}^{(2)} = [Y_{\ell,t+1}, \cdots , Y_{\ell,n}]$,

$$\underset{\sim}{Y}_{\ell,n} = \begin{bmatrix} Y_{\ell}^{(1)} \\ Y_{\ell}^{(2)} \end{bmatrix};$$

let $\underset{\sim}{u}_{\ell}^{(1)} = [u_{\ell,1}, \cdots , u_{\ell,t}]'$, $\underset{\sim}{u}_{\ell}^{(2)} = [u_{\ell,t+1}, \cdots , u_{\ell,n}]'$, $u_{\ell,n} = \begin{bmatrix} \underset{\sim}{u}_{\ell}^{(1)} \\ \underset{\sim}{u}_{\ell}^{(2)} \end{bmatrix}$,

$$\boldsymbol{\Theta}_{\ell} = [\boldsymbol{\theta}'_{\ell,0}, \ \boldsymbol{\theta}'_{\ell,1}, \cdots , \boldsymbol{\theta}'_{\ell,n}]', \ \boldsymbol{\phi} = [\boldsymbol{\theta}'_{1,0}, \ \mathbf{w}'_{1,1}, \cdots , \mathbf{w}'_{1,n}]'.$$

Let $\bar{\underset{\sim}{y}}_{\ell\text{-}1}$ denote all prior information (including the data) before observing any data from series Y_{ℓ}.

For model (3.3) consider the two scenarios described below.

Scenario A: Let $(\boldsymbol{\phi} \mid \bar{\underset{\sim}{y}}_0) \sim N(\boldsymbol{\mu}, \ \Sigma)$, $\underset{\sim}{u}_{1,n} \sim N(\mathbf{0}, \ \mathbf{U}_{1,n})$ sequences $\{u_{1,t}\}$, $\{\mathbf{w}_{1,t}\}$ are independent, $\boldsymbol{\mu}$, Σ, $\mathbf{U}_{1,n}$ are known; assume that after filtering ℓ-1 series,

$\ell=2, 3, \ldots, (\Theta_{\ell-1} \mid \bar{y}_{\ell-1}) \sim N(\mathbf{m}_{\ell-1}, \mathbf{C}_{\ell-1})$ has been obtained and that $\underline{u}_{\ell,n} \sim N(\mathbf{0}, \mathbf{U}_{\ell,n})$, $\mathbf{w}_{\ell,n} \sim N(\mathbf{0}, \mathbf{W}_{\ell,n})$, sequences $\{u_{\ell,t}\}$, $\{\mathbf{w}_{\ell,t}\}$ and $\{\Theta_{\ell-1,t} \mid \bar{y}_{\ell-1}\}$ are mutually independent, $(\Theta_{\ell,0} \mid \bar{y}_{\ell-1}) = (\Theta_{\ell-1,0} \mid \bar{y}_{\ell+1}) + \mathbf{s}_{\ell,0}$ with $\mathbf{s}_{\ell,0}$ known, $\mathbf{U}_{\ell,n}$ and $\mathbf{W}_{\ell,n}$ are also known.

Scenario B: Let $(\boldsymbol{\phi} \mid \bar{y}_0, U) \sim N(\boldsymbol{\mu}, U\boldsymbol{\Sigma})$, $(\underline{u}_{1,n} \mid U) \sim N(\mathbf{0}, U\mathbf{U}_{1,n})$, given U sequences $\{u_{1,t}\}$, $\{\mathbf{w}_{1,t}\}$ are independent; $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\mathbf{U}_{1,n}$ are known, U is unknown but $(U \mid \bar{y}_0) \sim IG\left(\frac{n_{1,0}}{2}, \frac{d_{1,0}}{2}\right)$. Assume that after filtering ℓ-1 series, $\ell=2, 3, \ldots, (\Theta_{\ell-1} \mid \bar{y}_{\ell-1}, U) \sim N(\mathbf{m}_{\ell-1}, U\mathbf{C}_{\ell-1})$ and $(U \mid \bar{y}_{\ell-1}) \sim IG\left(\frac{n_{\ell,0}}{2}, \frac{d_{\ell,0}}{2}\right)$ have been obtained; we also assume that $(\underline{u}_{\ell,n} \mid U) \sim N(\mathbf{0}, U\mathbf{U}_{\ell,n})$, $(\mathbf{w}_{\ell,n} \mid U) \sim N(0, U\mathbf{W}_{\ell,n})$, given U, sequences $\{u_{\ell,t}\}$, $\{\mathbf{w}_{\ell,t}\}$ and $\{\boldsymbol{\theta}_{\ell-1,t} \mid \bar{y}_{\ell-1}\}$ are mutually independent, $\mathbf{U}_{\ell,n}$, $\mathbf{W}_{\ell,n}$ are known, $(\boldsymbol{\theta}_{\ell,0} \mid \bar{y}_{\ell-1}, \cdot) = (\boldsymbol{\theta}_{\ell-1,0} \mid \bar{y}_{\ell-1}, \cdot) + \mathbf{s}_{\ell,0}$ with $\mathbf{s}_{\ell,0}$ known.

We have the following formulae for inference and forecasting.

1) For Scenario A

The inference and forecasting formulae of series 1 are

$$\text{i)} \quad \left[\begin{array}{c|c} \Theta_1 & \\ Y_1^{(1)} & \bar{y}_0 \\ Y_1^{(2)} & \end{array}\right] \sim N\left[\begin{bmatrix} \mathbf{m}_{\theta_1} \\ \mathbf{m}_{y_{1.1}} \\ \mathbf{m}_{y_{1.2}} \end{bmatrix}, \begin{bmatrix} \mathbf{C}_{\theta_1} & \mathbf{C}_{\theta_1 y_{1.1}} & \mathbf{C}_{\theta_1 y_{1.2}} \\ \mathbf{C}_{y_{1.1}\theta_1} & \mathbf{C}_{y_{1.1}} & \mathbf{C}_{y_{1.12}} \\ \mathbf{C}_{y_{1.2}\theta_1} & \mathbf{C}_{y_{1.21}} & \mathbf{C}_{y_{1.2}} \end{bmatrix}\right], \tag{3.5}$$

where $\quad \mathbf{m}_{\theta_1} = \mathbf{L}\,\boldsymbol{\mu}, \begin{pmatrix} \mathbf{m}_{y_{1.1}} \\ \mathbf{m}_{y_{1.2}} \end{pmatrix} = \mathbf{F}'\mathbf{m}_{\theta_1},$

$$C_{\theta_1} = L\,\Sigma\,L', \quad \begin{pmatrix} C_{y_{1.1}} & C_{y_{1.12}} \\ C_{y_{1.21}} & C_{y_{1.2}} \end{pmatrix} = F'C_{\theta_1}F + U_{1,n},$$

$$\begin{pmatrix} C_{\theta_1 y_{1.1}} & C_{\theta_1 y_{1.2}} \end{pmatrix} = \begin{bmatrix} C_{y_{1.1}\theta_1} \\ C_{y_{1.2}\theta_1} \end{bmatrix}' = C_{\theta_1}\,F,$$

$$L = \begin{bmatrix} I & & & & \\ G_1 & I & & & \\ G_2G_1 & G_2 & I & & \\ \vdots & \vdots & \vdots & \ddots & \\ G_n \cdots G_1 & G_n \cdots G_2 & G_n \cdots G_3 & \cdots & I \end{bmatrix},$$

$$F' = \begin{bmatrix} 0 & F_1' & & & \\ & 0 & F_2' & & \\ & & 0 & F_2' & \\ & & \ddots & \ddots & \\ & & & 0 & F_n' \end{bmatrix};$$

$$\text{ii)} \quad \left(\begin{matrix} \Theta \\ \underset{\sim}{Y}_1^{(2)} \end{matrix} \,\middle|\, \bar{y}_0, \underset{\sim}{y}_{1,t} \right) \sim N\left(\begin{pmatrix} \tilde{m}_{\theta_1} \\ \tilde{m}_{y_{1.2}} \end{pmatrix}, \begin{pmatrix} \tilde{C}_{\theta_1} & \tilde{C}_{\theta_1 y_{1.2}} \\ \tilde{C}_{y_{1.2}\theta_1} & \tilde{C}_{y_{1.2}} \end{pmatrix} \right), \tag{3.6}$$

where

$$\tilde{m}_{\theta_1} = m_{\theta_1} + C_{\theta_1 y_{1.1}} (C_{y_{1.1}})^{-1} (\underset{\sim}{y}_{1,t} - m_{y_{1.1}}),$$

$$\tilde{m}_{y_{1.2}} = m_{y_{1.2}} + C_{y_{1.21}} (C_{y_{1.1}})^{-1} (\underset{\sim}{y}_{1,t} - m_{y_{1.1}}),$$

$$\begin{pmatrix} \tilde{C}_{\theta_1} & \tilde{C}_{\theta_1 y_{1.2}} \\ \tilde{C}_{y_{1.2}\theta_1} & \tilde{C}_{y_{1.2}} \end{pmatrix} = \begin{pmatrix} C_{\theta_1} & C_{\theta_1 y_{1.2}} \\ C_{y_{1.2}\theta_1} & C_{y_{1.2}} \end{pmatrix} - \begin{pmatrix} C_{\theta_1 y_{1.1}} \\ C_{y_{1.21}} \end{pmatrix} C_{y_{1.1}}^{-1} \begin{pmatrix} C_{y_{1.1}\theta_1} & C_{y_{1.12}} \end{pmatrix};$$

iii) specifically, the k-step ahead predictive distribution is

$$(Y_{1,t+k} \mid \bar{y}_0, \underset{\sim}{y}_{1,t}) \sim N(\tilde{m}_{t+k}, \tilde{c}_{t+k,\, t+k}) \qquad \text{for } k=1, \dots, n-t, \tag{3.7}$$

where $\tilde{m}_{t+k}$ is the k-th component of $\tilde{\mathbf{m}}_{y_{1.2}}$,

$\tilde{c}_{t+k,\, t+k}$ is the k-th diagonal entry of $\tilde{C}_{y_{1.2}}$.

Recursively, for series ℓ, $\ell=2, 3, \dots$, we have

$$\text{i)} \quad \left[\begin{array}{c} \boldsymbol{\Theta}_\ell \\ \underset{\sim}{Y}_\ell^{(1)} \\ \underset{\sim}{Y}_\ell^{(2)} \end{array} \middle| \, \bar{y}_{\ell-1} \right] \sim N \left[\begin{bmatrix} \mathbf{m}_{\theta_\ell} \\ \mathbf{m}_{y_{\ell.1}} \\ \mathbf{m}_{y_{\ell,2}} \end{bmatrix}, \begin{bmatrix} C_{\theta_\ell} & C_{\theta_\ell y_{\ell.1}} & C_{\theta_\ell y_{\ell.2}} \\ C_{y_{\ell.1}\theta_\ell} & C_{y_{\ell.1}} & C_{y_{\ell.12}} \\ C_{y_{\ell.2}\theta_\ell} & C_{y_{\ell.21}} & C_{y_{\ell.2}} \end{bmatrix} \right], \tag{3.8}$$

where
$$\mathbf{m}_{\theta_\ell} = \mathbf{L}_\ell \ \mathbf{m}_{\ell-1} + \mathbf{s}_\ell, \quad \begin{pmatrix} \mathbf{m}_{y_{\ell.1}} \\ \mathbf{m}_{y_{\ell.2}} \end{pmatrix} = \mathbf{F}' \ \mathbf{m}_{\theta_\ell},$$

$$C_{\theta_\ell} = \mathbf{L}_\ell \, C_{\ell-1} \, \mathbf{L}_\ell' + \mathbf{M}_\ell \begin{bmatrix} \mathbf{0} & \\ & \mathbf{W}_{\ell,n} \end{bmatrix} \mathbf{M}_\ell',$$

$$\begin{pmatrix} C_{y_{\ell.1}\theta_\ell} \\ C_{y_{\ell.2}\theta_\ell} \end{pmatrix} = \begin{pmatrix} C_{\theta_\ell y_{\ell.1}} & C_{\theta_\ell y_{\ell.2}} \end{pmatrix}' = \mathbf{F}' C_{\theta_\ell},$$

$$\begin{pmatrix} C_{y_{\ell.1}} & C_{y_{\ell.12}} \\ C_{y_{\ell.21}} & C_{y_{\ell.2}} \end{pmatrix} = \mathbf{F}'\mathbf{C}_{\theta_\ell}\,\mathbf{F} + \mathbf{U}_{\ell,n}\,,$$

$$\mathbf{L}_\ell = \begin{bmatrix} \mathbf{I} & & & & \\ \mathbf{H}_{\ell,1} & \mathbf{J}_{\ell,1} & & & \\ \mathbf{H}_{\ell,2}\,\mathbf{H}_{\ell,1} & \mathbf{H}_{\ell,2}\mathbf{J}_{\ell,1} & \mathbf{J}_{\ell,2} & & \\ \vdots & \vdots & \vdots & \ddots & \\ \mathbf{H}_{\ell,n}\cdots\mathbf{H}_{\ell,1} & \mathbf{H}_{\ell,\,n}\cdots\mathbf{H}_{\ell,2}\mathbf{J}_{\ell,1} & \mathbf{H}_{\ell,n}\cdots\mathbf{H}_{\ell,3}\mathbf{J}_{\ell,2} & \cdots & \mathbf{J}_{\ell,n} \end{bmatrix},$$

$\mathbf{F}$ is specified as the same as in (3.5),

$$\mathbf{s}_\ell = \begin{bmatrix} \mathbf{I} \\ \mathbf{H}_{\ell,1} \\ \mathbf{H}_{\ell,2}\;\mathbf{H}_{\ell,1} \\ \vdots \\ \mathbf{H}_{\ell,n}\cdots\mathbf{H}_{\ell,1} \end{bmatrix} \mathbf{s}_{\ell,0}\,;$$

$$\mathbf{M}_\ell = \begin{bmatrix} \mathbf{I} & & & & \\ \mathbf{H}_{\ell,1} & \mathbf{I} & & & \\ \mathbf{H}_{\ell,2}\,\mathbf{H}_{\ell,1} & \mathbf{H}_{\ell,2} & \mathbf{I} & & \\ \vdots & \vdots & \vdots & \ddots & \\ \mathbf{H}_{\ell,n}\cdots\mathbf{H}_{\ell,1} & \mathbf{H}_{\ell,\,n}\cdots\mathbf{H}_{\ell,2} & \mathbf{H}_{\ell,n}\cdots\mathbf{H}_{\ell,3} & \cdots & \mathbf{I} \end{bmatrix};$$

$$\text{ii)}\quad \left(\begin{matrix} \mathbf{\Theta}_\ell \\ \underset{\sim}{Y}_\ell^{(2)} \end{matrix} \,\middle|\, \bar{\underset{\sim}{y}}_{\ell\text{-}1},\, \underset{\sim}{y}_{\ell,t}\right) \sim N\left(\begin{pmatrix} \tilde{\mathbf{m}}_{\theta_\ell} \\ \tilde{\mathbf{m}}_{y_{\ell.2}} \end{pmatrix},\ \begin{pmatrix} \tilde{C}_{\theta_\ell} & \tilde{C}_{\theta_\ell y_{\ell.2}} \\ \tilde{C}_{y_{\ell.2}\theta_\ell} & C_{y_{\ell.2}} \end{pmatrix}\right), \tag{3.9}$$

where
$$\tilde{\mathbf{m}}_{\theta_\ell} = \mathbf{m}_{\theta_\ell} + \mathbf{C}_{\theta_\ell y_{\ell.1}}\,(\mathbf{C}_{y_{\ell.1}})^{-1}\,(\underset{\sim}{y}_{\ell,t} - \mathbf{m}_{y_{\ell.1}}),$$

$$\tilde{m}_{y_{\ell.2}} = m_{y_{\ell.2}} + C_{y_{\ell.21}} (C_{y_{\ell.1}})^{-1} (\underset{\sim}{y}_{\ell,t} - m_{y_{\ell.1}}),$$

$$\begin{pmatrix} \tilde{C}_{\theta_\ell} & \tilde{C}_{\theta_\ell y_{\ell.2}} \\ \tilde{C}_{y_{\ell.2}\theta_\ell} & \tilde{C}_{y_{\ell.2}} \end{pmatrix} = \begin{pmatrix} C_{\theta_\ell} & C_{\theta_\ell y_{\ell.2}} \\ C_{y_{\ell.2}\theta_\ell} & C_{y_{\ell.2}} \end{pmatrix} - \begin{pmatrix} C_{\theta_\ell y_{\ell.1}} \\ C_{y_{\ell.21}} \end{pmatrix} C_{y_{\ell.1}}^{-1} \left(C_{y_{\ell.1}\theta_\ell} \;\; C_{y_{\ell.12}} \right);$$

iii) the k-step ahead predictive distribution is

$$(\underset{\sim}{Y}'_{\ell,t+k} \,|\, \tilde{Y}'_{\ell-1}, \underset{\sim}{y}_{\ell,t}) \sim N(\tilde{m}_{t+k}\,, \tilde{c}_{t+k,\, t+k}) \quad \text{for } k=1, 2, \ldots, n-t, \tag{3.10}$$

where $\tilde{m}_{t+k}$ is the k-th component of $\tilde{m}_{y_{\ell.2}}$,

$\tilde{c}_{t+k,\, t+k}$ is the k-th diagonal entry of $\tilde{C}_{y_{\ell.2}}$.

2) For Scenario B

The inference and forecasting formulae of series 1 are

$$\text{i)} \left[\begin{array}{c|c} \boldsymbol{\Theta}_1 & \\ \underset{\sim}{Y}_1^{(1)} & \tilde{y}_0, U \\ \underset{\sim}{Y}_1^{(2)} & \end{array} \right] \sim N \left[\begin{bmatrix} m_{\theta_1} \\ m_{y_{1.1}} \\ m_{y_{1.2}} \end{bmatrix}, U \begin{bmatrix} C_{\theta_1} & C_{\theta_1 y_{1.1}} & C_{\theta_1 y_{1.2}} \\ C_{y_{1.1}\theta_1} & C_{y_{1.1}} & C_{y_{1.12}} \\ C_{y_{1.2}\theta_1} & C_{y_{1.21}} & C_{y_{1.2}} \end{bmatrix} \right], \tag{3.11a}$$

$$\left[\begin{array}{c|c}\boldsymbol{\Theta}_1 & \\ \underset{\sim}{Y}_1^{(1)} & \bar{\underset{\sim}{y}}_0 \\ \underset{\sim}{Y}_1^{(2)} & \end{array}\right] \sim T_{n_{1,0}}\left[\begin{bmatrix}\mathbf{m}_{\theta_1} \\ \mathbf{m}_{y_{1.1}} \\ \mathbf{m}_{y_{1.2}}\end{bmatrix}, \ \frac{d_{1,0}}{n_{1,0}}\begin{bmatrix}C_{\theta_1} & C_{\theta_1 y_{1.1}} & C_{\theta_1 y_{1.2}} \\ C_{y_{1.1}\theta_1} & C_{y_{1.1}} & C_{y_{1.12}} \\ C_{y_{1.2}\theta_1} & C_{y_{1.21}} & C_{y_{1.2}}\end{bmatrix}\right], \quad (3.11b)$$

where all submatrices are specified as the same as in (3.5);

$$\text{ii)} \quad (U \mid \bar{\underset{\sim}{y}}_0, \underset{\sim}{y}_{1,t}) \sim IG\left(\frac{n_{1,t}}{2}, \frac{d_{1,t}}{2}\right) \qquad (3.12a)$$

$$\left(\begin{matrix}\boldsymbol{\Theta}_1 \\ \underset{\sim}{Y}_1^{(2)}\end{matrix} \mid \bar{\underset{\sim}{y}}_0, \underset{\sim}{y}_{1,t}, U\right) \sim N\left(\left(\begin{matrix}\tilde{\mathbf{m}}_{\theta_1} \\ \tilde{\mathbf{m}}_{y_{1.2}}\end{matrix}\right), \ U\left(\begin{matrix}\tilde{C}_{\theta_1} & \tilde{C}_{\theta_1 y_{1.2}} \\ \tilde{C}_{y_{1.2}\theta_1} & \tilde{C}_{y_{1.2}}\end{matrix}\right)\right), \qquad (3.12b)$$

$$\left(\begin{matrix}\boldsymbol{\Theta}_1 \\ \underset{\sim}{Y}_1^{(2)}\end{matrix} \mid \bar{\underset{\sim}{y}}_0, \underset{\sim}{y}_{1,t}\right) \sim T_{n_{1,t}}\left(\left(\begin{matrix}\tilde{\mathbf{m}}_{\theta_1} \\ \tilde{\mathbf{m}}_{y_{1.2}}\end{matrix}\right), \ \frac{d_{1,t}}{n_{1,t}}\left(\begin{matrix}\tilde{C}_{\theta_1} & \tilde{C}_{\theta_1 y_{1.2}} \\ \tilde{C}_{y_{1.2}\theta_1} & \tilde{C}_{y_{1.2}}\end{matrix}\right)\right), \qquad (3.12c)$$

where $\quad n_{1,t} = n_{1,0} + t, \qquad d_{1,t} = d_{1,0} + d'_{1,t},$

$$d'_{1,t} = (\underset{\sim}{y}_{1,t} - \mathbf{m}_{y_{1.1}})' (C_{y_{1.1}})^{-1} (\underset{\sim}{y}_{1,t} - \mathbf{m}_{y_{1.1}})$$

all submatrices are specified as the same as in (3.6);

iii) specifically, the k=step ahead predictive distribution is

$$(Y_{1,t+k} \mid \bar{\underset{\sim}{y}}_0, \underset{\sim}{y}_1^{(1)}) \sim T_{n_{1,0}}(\tilde{m}_{t+k}, \ \frac{d_{1,t}}{n_{1,t}} \ \tilde{c}_{t+k,\, t+k}), \quad \text{for } k=1, \dots, n-t, \qquad (3.13)$$

where $\quad \tilde{m}_{t+k}^{(t)}$ is the k-th component of $\tilde{\mathbf{m}}_{y_{1.2}}$,

$\tilde{c}_{t+k,\ t+k}$ is the k-th diagonal entry of $\tilde{C}_{y_{1.2}}$.

Recursively, for series ℓ, ℓ=2, 3, ... , we have

$$\text{i)}\quad \left[\begin{array}{c|c} \boldsymbol{\Theta}_\ell & \\ \underset{\sim}{Y}_\ell^{(1)} & \bar{\underset{\sim}{y}}_{\ell\text{-}1}, U \\ \underset{\sim}{Y}_\ell^{(2)} & \end{array}\right] \sim N\left[\begin{bmatrix} m_{\theta_\ell} \\ m_{y_{\ell.1}} \\ m_{y_{\ell.2}} \end{bmatrix}, U\begin{bmatrix} C_{\theta_\ell} & C_{\theta_\ell y_{\ell.1}} & C_{\theta_\ell y_{\ell.2}} \\ C_{y_{\ell.1}\theta_\ell} & C_{y_{\ell.1}} & C_{y_{\ell.12}} \\ C_{y_{\ell.2}\theta_\ell} & C_{y_{\ell.21}} & C_{y_{\ell.2}} \end{bmatrix}\right], \tag{3.14a}$$

$$\left[\begin{array}{c|c} \boldsymbol{\Theta}_\ell & \\ \underset{\sim}{Y}_\ell^{(1)} & \bar{\underset{\sim}{y}}_{\ell\text{-}1} \\ \underset{\sim}{Y}_\ell^{(2)} & \end{array}\right] \sim T_{n_{\ell,0}}\left[\begin{bmatrix} m_{\theta_\ell} \\ m_{y_{\ell.1}} \\ m_{y_{\ell.2}} \end{bmatrix}, \frac{d_{\ell,0}}{n_{\ell,0}}\begin{bmatrix} C_{\theta_\ell} & C_{\theta_\ell y_{\ell.1}} & C_{\theta_\ell y_{\ell.2}} \\ C_{y_{\ell.1}\theta_\ell} & C_{y_{\ell.1}} & C_{y_{\ell.12}} \\ C_{y_{\ell.2}\theta_\ell} & C_{y_{\ell.21}} & C_{y_{\ell.2}} \end{bmatrix}\right], \tag{3.14b}$$

where all submatrices are specified as the same as in (3.8);

$$\text{ii)}\quad (U \mid \bar{\underset{\sim}{y}}_{\ell\text{-}1}, \underset{\sim}{y}_{\ell,t}) \sim IG\left(\frac{n_{\ell,t}}{2}, \frac{d_{\ell,t}}{2}\right) \tag{3.15a}$$

$$\left(\begin{matrix} \boldsymbol{\Theta}_\ell \\ \underset{\sim}{Y}_\ell^{(2)} \end{matrix} \mid \bar{\underset{\sim}{y}}_{\ell\text{-}1}, \underset{\sim}{y}_{\ell,t}, U\right) \sim N\left(\begin{pmatrix} \tilde{m}_{\theta_\ell} \\ \tilde{m}_{y_{\ell.2}} \end{pmatrix}, U \begin{pmatrix} \tilde{C}_{\theta_\ell} & \tilde{C}_{\theta_\ell y_{\ell.2}} \\ \tilde{C}_{y_{\ell.2}\theta_\ell} & \tilde{C}_{y_{\ell.2}} \end{pmatrix}\right), \tag{3.15b}$$

$$\left(\begin{matrix} \boldsymbol{\Theta}_\ell \\ \underset{\sim}{Y}_\ell^{(2)} \end{matrix} \mid \bar{\underset{\sim}{y}}_{\ell\text{-}1}, \underset{\sim}{y}_{\ell,t}\right) \sim T_{n_{\ell,t}}\left(\begin{pmatrix} \tilde{m}_{\theta_\ell} \\ \tilde{m}_{y_{\ell.2}} \end{pmatrix}, \frac{d_{\ell,t}}{n_{\ell,t}} \begin{pmatrix} \tilde{C}_{\theta_\ell} & \tilde{C}_{\theta_\ell y_{\ell.2}} \\ \tilde{C}_{y_{\ell.2}\theta_\ell} & \tilde{C}_{y_{\ell.2}} \end{pmatrix}\right), \tag{3.15c}$$

where $\quad n_{\ell,t} = n_{\ell,0} + t, \qquad d_{\ell,t} = d_{\ell,0} + d'_{\ell,t}$,

$$d'_{\ell,t} = (\underset{\sim}{y}_{\ell,t} - m_{y_{\ell.1}})' (C_{y_{\ell.1}})^{-1} (\underset{\sim}{y}_{\ell,t} - m_{y_{\ell.1}}),$$

all submatrices are specified as the same as in (3.9);

iii) the k=step ahead predictive distribution is

$$(Y_{\ell,t+k} \mid \tilde{\underset{\sim}{y}}_{\ell-1}, \underset{\sim}{y}_{\ell}^{(1)}) \sim T_{n_{\ell,t}} (\tilde{m}_{t+k}, \frac{d_{\ell,t}}{n_{\ell,t}} \tilde{c}_{t+k,\, t+k}), \text{ for } k=1, 2, \dots, n\text{-}t, \qquad (3.16)$$

where $\tilde{m}_{t+k}$, $\tilde{c}_{t+k,\, t+k}$ are specified as the same as in (3.10).

It can be shown that equations (2.2) - (2.5) are merely the marginal distribution of (3.6) and (3.12c) respectively. Since the DLM with leading indicators deals with the series of interest in conjunction with indicator series, any assumption on the independence of $\boldsymbol{\theta}_{\ell,t-1}$ and $\boldsymbol{\theta}_{\ell,t}$ is irrelevant. The only way to treat the correlations among all state parameters, is to represent the probabilistic structure of the DLM or the DLM with leading indicators, and the mechanism of the filtering procedure, in terms of the joint distribution of the state parameter space.

4. APPLICATIONS

We apply our approach to the data given in Section 1.

If the sequence $[y_1, y_2, \dots, y_t, y_{t+1}, \dots, y_m]$ is observed from the time series of interest and we were to use the first t observations to predict the next m-t values , then we specify the cumulative square error (CSE) as

$$\text{CSE}(m, t) = \sum_{k=1}^{m-t} (y_{t+k} - E(\hat{Y}_{t+k} \mid \underset{\sim}{y}_t))^2 ,$$

where $(\hat{Y}_{t+k} \mid \underset{\sim}{y}_t)$ is the k-step ahead prediction based on observations $\underset{\sim}{y}_t$. We use CSE to measure

the forecasting performance of our approach. The smaller the CSE(m, t) is, the better the forecasts will be since CSE(m, t) reflects the total loss of the (m-t) forecasts when the loss function is a quadratic function.

We select an *S-shaped growth model*, a *segmented linear growth model* and a *linear growth model* to analyze Data Sets One, Two and Three respectively. We set the conditions for data analysis in the scheme described as scenario B. Specifically, for the DLM, we specify the initial conditions as

the prior distribution of U is an IG (2.5, 1), i.e. $n_0 = 5$, $d_0 = 2$;

$(v_t \mid U) \sim N(0, 0.01\,U)$; $\quad (w_t \mid U) \sim N(0, 0.2\,U)$;

$(\theta_0 \mid \underset{\sim}{y}_0, U) \sim N(Y_0, 0.01\,U)$; $(\beta_0 \mid \underset{\sim}{y}_0, U) \sim N(y_1 - Y_0, 0.1\,U)$;

Y_0 is assumed available; $\quad (\theta_0 \mid \underset{\sim}{y}_0, U)$ is independent of $(\beta_0 \mid \underset{\sim}{y}_0, U)$.

Analogous to the above specification, the initial conditions for the DLM with leading indicators are

$$\left(\begin{matrix}\theta_{\ell,0}\\ \beta_{\ell,0}\end{matrix} \,\middle|\, \underset{\sim}{\bar{y}}_{\ell\text{-}1}, U\right) \sim N\left(\left(\begin{matrix}Y_{\ell,0}\\ y_{\ell,1}-Y_{\ell,0}\end{matrix}\right), \; U\textstyle\sum_0\right), \quad \text{for} \quad \ell=2, 3, \ldots$$

$$\text{with } \textstyle\sum_0 = E\left(\left(\begin{matrix}\theta_{\ell\text{-}1,0} - E(\theta_{\ell\text{-}1,0})\\ \beta_{\ell\text{-}1,0} - E(\beta_{\ell\text{-}1,0})\end{matrix}\right)\left(\theta_{\ell\text{-}1,0} - E(\theta_{\ell\text{-}1,0}) \quad \beta_{\ell\text{-}1,0} - E(\beta_{\ell\text{-}1,0})\right) \,\middle|\, \underset{\sim}{\bar{y}}_{\ell\text{-}1}, U\right)/U\ ;$$

$$\left(\begin{matrix}\theta_{1,0}\\ \beta_{1,0}\end{matrix} \,\middle|\, \underset{\sim}{y}_{1,0}, U\right) \sim N\left(\left(\begin{matrix}Y_{1,0}\\ y_{1,1}-Y_{1,0}\end{matrix}\right); \; U\left(\begin{matrix}0.01 & 0\\ 0 & 0.1\end{matrix}\right)\right), \quad (U \mid \underset{\sim}{y}_{1,0}) \sim IG(2.5, 1);$$

$$(v_{\ell,t} \mid U) \sim N(0, 0.01U); \qquad (w_{\ell,t} \mid U) \sim N(0, 0.2U).$$

Part of our analysis results are listed in Tables 4.1 - 4.4. These results show that the use of DLM with leading indicator can improve forecasting significantly, incorporation of two or more

indicator series may not be more effective than incorporation of one indicator series (cf. Table 4.2). Table 4.4 offers an example to assign the value of γ intuitively. From Figure 1.3 the data plot of series Y_3 looks quite similar to that of series Y_2. So a small γ (namely, $\gamma = 0.05$) leads to much better forecasts than a bigger one ($\gamma = 0.975$ here). It also gives us an idea of how the variances of the predictions will be if the assignment of d_0 is not appropriate. In this case $d_0 = 2 \gg d'_t = 0.005684$ (when $\gamma = 0.975$), $E(Y_{40} | \underset{\sim}{y}_{10}) = 2.4977$, $V(Y_{40} | \underset{\sim}{y}_{10}) = 63.0757$. If we shrink d_0 from 2 to 0.05, the above variance will reduce to its 1/40 of the current values and the mean remains unchanged.

Table 4.1 A Comparison of Forecasts of Y_3 of Data Set One Series from the Growth Model and from the 2-d Model ($\gamma = .05$)

t	Observation	Forecasts from S-shaped Growth Model		Forecasts from the DLM with leading indicators (Indicator series: Y_1)	
		mean	variance	mean	variance
21	293.7	292.51	11.65	293.98	5.25
22	300.9	299.71	18.77	301.44	6.27
23	307.7	305.83	29.40	307.45	7.19
24	298.0	311.04	43.59	312.15	8.02
25	317.6	315.46	61.19	315.70	8.78
26	319.9	319.22	81.99	318.30	9.51
27	320.9	322.42	105.72	320.16	10.21
28	321.8	325.13	132.09	321.51	10.92
29	322.7	327.44	160.80	322.54	11.62
30	323.3	329.41	191.59	323.34	12.32
31	323.9	331.07	224.21	323.91	13.03
32	324.1	332.49	258.42	324.29	13.74
33	324.2	333.70	294.02	324.53	14.45
34	324.4	334.72	330.80	324.7	15.17
35	324.5	335.59	368.62	324.82	15.88
36	324.6	336.96	407.32	324.92	16.59
37	324.7	337.50	446.79	325.02	17.31
38	324.7	337.95	486.91	325.11	18.16
39	324.7	338.34	527.59	325.18	19.60
	CSE	1313.48		208.408	
	d'_{20}	111.567		39.006	

Table 4.2 A Comparison of Forecasts of Y_6 of Data Set One from the Growth Model and from the 2-d Model (with γ=.05)

t	Observation	Forecasts from S-shaped Growth model		Forecasts from the DLM with leading indicators			
				Indicator series: Y_5		Indicator series: Y_4 and Y_5	
		mean	variance	mean	variance	mean	variance
21	314.4	316.13	38.80	317.96	11.95	317.62	14.84
22	322.8	325.57	62.49	327.67	14.31	327.11	17.61
23	330.2	333.60	97.87	335.46	16.47	334.76	20.02
24	337.6	340.43	145.12	341.55	18.42	340.85	22.20
25	341.7	346.23	203.73	346.08	20.20	345.39	24.27
26	345.8	351.17	272.99	349.27	21.89	348.52	26.29
27	346.5	355.35	351.98	351.47	23.54	350.55	28.30
28	348.4	358.91	439.79	353.14	25.17	352.30	30.31
29	350.2	361.94	535.40	354.39	26.81	353.42	32.33
30	351.7	364.51	637.91	355.63	28.45	354.65	34.34
31	353.3	366.70	746.51	359.97	30.10	356.08	36.35
32	354.8	368.56	860.41	358.22	31.76	357.32	38.36
33	356.4	370.15	978.94	359.28	33.42	358.28	40.37
CSE		1313.48		224.58		149.56	
d'_{20}		376.121		319.94		264.66	

Table 4.3 A Comparison of Forecasts of Y_3 of Data Set Two from the Growth Model and from the 2-d Model ($\gamma = .975$)

t	Observation	Forecasts from the Segmented Linear Growth Model		Forecasts from the DLM with leading indicators (Indicator series: Y_2)	
		mean	variance	mean	variance
11	81.7	82.49	7.49	82.75	6.72
12	86.5	87.15	13.50	87.58	11.87
13	89.01	91.34	22.65	92.42	20.48
14	91.87	95.53	36.50	97.24	33.32
15	94.27	99.72	56.20	102.03	51.08
16	96.71	103.91	82.88	106.80	74.41
17	99.45	108.10	117.68	111.53	103.90
18	101.8	112.29	161.75	116.25	140.10
19	104.2	116.48	216.22	120.94	183.49
CSE		437.366		839.528	
d'_{10}		35.024		45.546	

Table 4.4 A Comparison of Forecasts of Y_3 of Data Set Three from the Growth Model and from the 2-d Model

t	Observation	Forecasts from the linear Growth Model: mean	Forecasts from the linear Growth Model: variance	Forecasts from the DLM with leading indicators (Indicator series: Y_2), $\gamma = .975$: mean	$\gamma = .975$: variance	$\gamma = .05$: mean	$\gamma = .05$: variance
11	.974	.9677	.4059	.9686	.1299	.9875	.0880
12	1.098	1.0210	.7314	1.0224	.2295	1.0462	.1051
13	1.162	1.0742	1.2883	1.0762	.3961	1.1009	.1210
14	1.196	1.1274	2.13821	1.1300	.6443	1.1527	.1353
15	1.196	1.1807	3.3428	1.1837	.9877	1.2027	.1484
CSE		.019		.018		.009	
d'_{10}		.005507		.005684		.002693	

5. FUTURE WORK

Our contribution in this effort on developing improved forecasting methods is the expansion of the DLM such that it can incorporate indicator series to improve forecasting. And because we have not added any additional restriction on the 2-d model, our methodology can work for general two-dimensional dynamic linear models provided they can be specified as sequential models of parametric structure. Thus, our approach has a wide field of applications, for instances, the two dimensional filter problems in image processing, and some estimation problems related to Markov random fields.

Despite the results that we have had so far, we need to do further work to more generalize our achievements. Namely,

1) Weight γ can be specified as a distribution if the forecaster is not sure to specify γ at a fixed value. This specification will make the use of the DLM with leading indicators much more difficult and complicated. However, we believe, with the help of some simulation techniques, say, Gibbs sampling, we can solve the forecasting problem of the DLM with leading indicators with weight γ specified as a distribution.

2) Explore new approaches to reduce the variances of the forecasts. From (3.12) and (3.15), we can see that the posterior mean of U - a main factor affecting the variances of the forecasts - is determined by the specification of the used DLM. So, another direction of our future work is to study new model monitoring techniques to reduce the variances of the forecasts.

APPENDIX

Derivation of Predictive Distributions

We derive here the inference and forecasting formulae given in Section 3, via some elementary distribution theory. Specifically, we need to apply the following well-known propositions pertaining to the properties of the multivariate normal distribution.

Proposition 1 If $\underset{\sim}{X} \sim N(\mu, \Sigma)$, then $C\underset{\sim}{X} \sim N(C\mu, C\Sigma C')$.

Proposition 2 If $\begin{pmatrix} \underset{\sim}{X}_1 \\ \underset{\sim}{X}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$, then

i) $\underset{\sim}{X}_i \sim N(\mu_i, \Sigma_{ii})$, $i=1, 2$;

ii) $(\underset{\sim}{X}_1 \mid \underset{\sim}{X}_2) \sim N(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underset{\sim}{X}_2 - \mu_2), \ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$, if Σ_{22} is nonsingular.

Proposition 3 If $\underset{\sim}{X}$ is a p-vector and $(\underset{\sim}{X} \mid \phi) \sim N(\mu, \phi^{-1}\Sigma)$, $\phi \sim G(\frac{n}{2}, \frac{d}{2})$ i.e. ϕ is a gamma random variable with shape parameter $\frac{n}{2}$ and scale parameter $\frac{d}{2}$, then

i) $(\phi \mid \underset{\sim}{X}=\underset{\sim}{x}) \sim G(\frac{n^*}{2}, \frac{d^*}{2})$ with $n^* = n+p$ and $d^* = d+(\underset{\sim}{x}-\mu)'\Sigma^{-1}(\underset{\sim}{x}-\mu)$;

ii) $\underset{\sim}{X} \sim T_n(\mu, \frac{d}{n}\Sigma)$.

A1. Derive the inference and forecasting formulae of Y_1.

Recall the model form for series 1

$$\begin{cases} y_{1,t} = F_t' \theta_{1,t} + u_{1,t} \\ \theta_{1,t} = G_t \theta_{1,t-1} + w_{1,t}. \end{cases} \tag{A.1}$$

From the second equation follows

$$\begin{aligned} \boldsymbol{\theta}_{1,t} &= G_t G_{t-1} \cdots G_1 \boldsymbol{\theta}_{1,0} + G_t \cdots G_2 \mathbf{w}_{1,1} + G_t \cdots G_3 \mathbf{w}_{1,2} + \cdots + \mathbf{w}_{1,t} \\ &= [G_t \cdots G_1, G_t \cdots G_2, \cdots, \mathbf{I}, \mathbf{0}, \cdots, \mathbf{0}]\, \boldsymbol{\phi}, \qquad \text{for } t=1, 2, \ldots, n, \text{ and} \\ \boldsymbol{\theta}_{1,0} &= [\mathbf{I}, \mathbf{0}, \ldots, \mathbf{0}]\, \boldsymbol{\phi}, \end{aligned}$$

where $\boldsymbol{\phi} = [\boldsymbol{\theta}_{1,0}' \ \mathbf{w}_{1,1}' \cdots \mathbf{w}_{1,n}']'$. Thus, $\boldsymbol{\Theta}_1 = [\boldsymbol{\theta}_{1,0}' \ \boldsymbol{\theta}_{1,1}' \cdots \boldsymbol{\theta}_{1,n}']' = \mathbf{L}\, \boldsymbol{\phi}$ with

$$\mathbf{L} = \begin{bmatrix} \mathbf{I} & & & & \\ G_1 & \mathbf{I} & & & \\ G_2 G_1 & G_2 & \mathbf{I} & & \\ \vdots & \vdots & \vdots & \ddots & \\ G_n \cdots G_1 & G_n \cdots G_2 & G_n \cdots G_3 & \cdots & \mathbf{I} \end{bmatrix}.$$

And from the first equation of (A.1) follows

$$Y_{1,t} = [\underbrace{\mathbf{0} \cdots \mathbf{0}}_{t\ \mathbf{0}\text{'s}} \ F_t' \ \mathbf{0} \cdots \mathbf{0}]\ \boldsymbol{\Theta}_1 + u_t \qquad \text{for } t=1, \ldots n\,.$$

Therefore $\underset{\sim}{Y}_{1,n} = \mathbf{F}'\,\boldsymbol{\Theta}_1 + \underset{\sim}{u}_{1,n} = \mathbf{F}'\mathbf{L}\,\boldsymbol{\phi} + \underset{\sim}{u}_{1,n}$ with

$$\mathbf{F}' = \begin{bmatrix} \mathbf{0} & \mathbf{F}_1' & & & & \\ & \mathbf{0} & \mathbf{F}_2' & & & \\ & & \mathbf{0} & \mathbf{F}_2' & & \\ & & & \ddots & \ddots & \\ & & & & \mathbf{0} & \mathbf{F}_n' \end{bmatrix} .$$

Under scenario A, i.e., $(\phi \mid \bar{y}) \sim N(\boldsymbol{\mu}, \Sigma)$, $\underset{\sim}{u}_{1,n} \sim N(\mathbf{0}, \mathbf{U}_{1,n})$, and $\{u_{1,t}\}$ $\{\mathbf{w}_{1,t}\}$ are independent sequences, applying proposition 1 yields

$$\left[\begin{array}{c|c|} \boldsymbol{\Theta}_1 & \\ \underset{\sim}{Y}_1^{(1)} & \bar{\underset{\sim}{y}}_0 \\ \underset{\sim}{Y}_1^{(2)} & \end{array}\right] \sim N\left[\begin{bmatrix} m_{\theta_1} \\ m_{y_{1.1}} \\ m_{y_{1.2}} \end{bmatrix}, \begin{bmatrix} C_{\theta_1} & C_{\theta_1 y_{1.1}} & C_{\theta_1 y_{1.2}} \\ C_{y_{1.1}\theta_1} & C_{y_{1.1}} & C_{y_{1.12}} \\ C_{y_{1.2}\theta_1} & C_{y_{1.21}} & C_{y_{1.2}} \end{bmatrix}\right],$$

where $\quad \mathbf{m}_{\theta_1} = \mathbf{L}\,\boldsymbol{\mu}, \qquad (\mathbf{m}_{y_{1.1}}' \;\mathbf{m}_{y_{1.2}}')' = \mathbf{F}'\,\mathbf{L}\,\boldsymbol{\mu}, \quad \mathbf{C}_{\theta_1 y_{1.1}} = \mathbf{C}_{y_{1.1}\theta_1}',$

$$\mathbf{C}_{y_{1.2}\theta_1} = \mathbf{C}_{\theta_1 y_{1.2}}', \; \mathbf{C}_{\theta_1} = \mathbf{L}\,\Sigma\,\mathbf{L}', \; (\mathbf{C}_{\theta_1 y_{1.1}} \; \mathbf{C}_{\theta_1 y_{1.2}}) = \mathbf{L}\,\Sigma\,\mathbf{L}'\mathbf{F},$$

$$\begin{pmatrix} C_{y_{1.1}} & C_{y_{1.12}} \\ C_{y_{1.21}} & C_{y_{1.2}} \end{pmatrix} = \mathbf{F}'\,\mathbf{L}\,\Sigma\,\mathbf{L}'\mathbf{F} + \mathbf{U}_{1,n} \,.$$

The above one is the prior joint distribution of $\boldsymbol{\Theta}_1$ and $\underset{\sim}{Y}_{1,n}$. Then, applying proposition 2 to the joint distribution, we obtain the conditional distribution of $\boldsymbol{\Theta}_1$ and $\underset{\sim}{Y}_1^{(2)}$ given observations $\underset{\sim}{y}_1^{(1)} = \underset{\sim}{y}_{1,t} = (y_{1,1}, \cdots, y_{1,t})'$ as

$$\begin{pmatrix} \boldsymbol{\Theta} \\ \underset{\sim}{Y}_1^{(2)} \end{pmatrix} \Big| \, \bar{\underset{\sim}{y}}_0, \underset{\sim}{y}_{1,t} \sim N\left(\begin{pmatrix} \tilde{\mathbf{m}}_{\theta_1} \\ \tilde{\mathbf{m}}_{y_{1.2}} \end{pmatrix}, \begin{pmatrix} \tilde{C}_{\theta_1} & \tilde{C}_{\theta_1 y_{1.2}} \\ \tilde{C}_{y_{1.2}\theta_1} & \tilde{C}_{y_{1.2}} \end{pmatrix}\right),$$

where

$$\begin{aligned} \tilde{\mathbf{m}}_{\theta_1} &= \mathbf{m}_{\theta_1} + C_{\theta_1 y_{1.1}} (C_{y_{1.1}})^{-1} (\underset{\sim}{y}_{1,t} - \mathbf{m}_{y_{1.1}}), \\ \tilde{\mathbf{m}}_{y_{1.2}} &= \mathbf{m}_{y_{1.2}} + C_{y_{1.21}} (C_{y_{1.1}})^{-1} (\underset{\sim}{y}_{1,t} - \mathbf{m}_{y_{1.1}}), \end{aligned}$$

$$\begin{pmatrix} \tilde{C}_{\theta_1} & \tilde{C}_{\theta_1 y_{1.2}} \\ \tilde{C}_{y_{1.2}\theta_1} & \tilde{C}_{y_{1.2}} \end{pmatrix} = \begin{pmatrix} C_{\theta_1} & C_{\theta_1 y_{1.2}} \\ C_{y_{1.2}\theta_1} & C_{y_{1.2}} \end{pmatrix} - \begin{pmatrix} C_{\theta_1 y_{1.1}} \\ C_{y_{1.21}} \end{pmatrix} C_{y_{1.1}}^{-1} \left(C_{y_{1.1}\theta_1} \;\; C_{y_{1.12}} \right).$$

In particular, the joint predictive distribution of $\underset{\sim}{Y}_1^{(2)}$ given observations $\underset{\sim}{y}_{1,t}$ is a multivariate normal specified by

$$(\underset{\sim}{Y}_1^{(2)} \,|\, \bar{\underset{\sim}{y}}_0, \underset{\sim}{y}_{1,t}) \sim N(\tilde{\mathbf{m}}_{y_{1.2}}, \tilde{C}_{y_{1.2}}).$$

Under scenario B, i.e. $(\boldsymbol{\phi} \,|\, \bar{\underset{\sim}{y}}_0, U) \sim N(\boldsymbol{\mu}, U\Sigma)$, $(\underset{\sim}{u}_{1,n} \,|\, U) \sim N(\mathbf{0}, UU_{1,n})$, given U sequences $\{u_{1,t}\}$ and $\{\mathbf{w}_{1,t}\}$ are independent, $(U \,|\, \bar{\underset{\sim}{y}}_0) \sim IG\left(\frac{n_{1,0}}{2}, \frac{d_{1,0}}{2}\right)$, then, similar to the above case, we have

$$\left[\begin{array}{c|c} \mathbf{\Theta}_1 & \\ \underset{\sim}{Y}_1^{(1)} & \bar{y}_0, U \\ \underset{\sim}{Y}_1^{(2)} & \end{array}\right] \sim N\left[\left[\begin{array}{c} \mathbf{m}_{\theta_1} \\ \mathbf{m}_{y_{1.1}} \\ \mathbf{m}_{y_{1.2}} \end{array}\right], U\left[\begin{array}{ccc} C_{\theta_1} & C_{\theta_1 y_{1.1}} & C_{\theta_1 y_{1.2}} \\ C_{y_{1.1}\theta_1} & C_{y_{1.1}} & C_{y_{1.12}} \\ C_{y_{1.2}\theta_1} & C_{y_{1.21}} & C_{y_{1.2}} \end{array}\right]\right],$$

$$\left[\begin{array}{c|c} \mathbf{\Theta}_1 & \\ \underset{\sim}{Y}_1^{(1)} & \bar{y}_0 \\ \underset{\sim}{Y}_1^{(2)} & \end{array}\right] \sim T_{n_{1,0}}\left[\left[\begin{array}{c} \mathbf{m}_{\theta_1} \\ \mathbf{m}_{y_{1.1}} \\ \mathbf{m}_{y_{1.2}} \end{array}\right], \frac{d_{1,0}}{n_{1,0}}\left[\begin{array}{ccc} C_{\theta_1} & C_{\theta_1 y_{1.1}} & C_{\theta_1 y_{1.2}} \\ C_{y_{1.1}\theta_1} & C_{y_{1.1}} & C_{y_{1.12}} \\ C_{y_{1.2}\theta_1} & C_{y_{1.21}} & C_{y_{1.2}} \end{array}\right]\right],$$

where all submatrices are specified as the same as under scenario A.

Applying proposition 3 to the distribution of $\left(\begin{array}{c} \mathbf{\Theta}_1 \\ \underset{\sim}{Y}_1^{(2)} \end{array} \mid U, \bar{y}_0, \underset{\sim}{y}_{1,t}\right)$, we obtain

$$(U \mid \bar{y}_0, \underset{\sim}{y}_{1,t}) \sim IG\left(\frac{n_{1,t}}{2}, \frac{d_{1,t}}{2}\right),$$

$$\left(\begin{array}{c} \mathbf{\Theta}_1 \\ \underset{\sim}{Y}_1^{(2)} \end{array} \mid \bar{y}_0, \underset{\sim}{y}_{1,t}, U\right) \sim N\left(\left(\begin{array}{c} \tilde{\mathbf{m}}_{\theta_1} \\ \tilde{\mathbf{m}}_{y_{1.2}} \end{array}\right), U\left(\begin{array}{cc} \tilde{C}_{\theta_1} & \tilde{C}_{\theta_1 y_{1.2}} \\ \tilde{C}_{y_{1.2}\theta_1} & \tilde{C}_{y_{1.2}} \end{array}\right)\right),$$

$$\left(\begin{array}{c} \mathbf{\Theta}_1 \\ \underset{\sim}{Y}_1^{(2)} \end{array} \mid \bar{y}_0, \underset{\sim}{y}_{1,t}\right) \sim T_{n_{1,t}}\left(\left(\begin{array}{c} \tilde{\mathbf{m}}_{\theta_1} \\ \tilde{\mathbf{m}}_{y_{1.2}} \end{array}\right), \frac{d_{1,t}}{n_{1,t}}\left(\begin{array}{cc} \tilde{C}_{\theta_1} & \tilde{C}_{\theta_1 y_{1.2}} \\ \tilde{C}_{y_{1.2}\theta_1} & \tilde{C}_{y_{1.2}} \end{array}\right)\right),$$

where $\quad n_{1,t} = n_{1,0} + t, \qquad d_{1,t} = d_{1,0} + d'_{1,t},$

$$d'_{1,t} = (\underset{\sim}{y}_{1,t} - \mathbf{m}_{y_{1.1}})' (C_{y_{1.1}})^{-1} (\underset{\sim}{y}_{1,t} - \mathbf{m}_{y_{1.1}}).$$

A2. Derive the inference and forecasting formulae of Y_ℓ.

The model formulating Y_ℓ is

$$\begin{cases} Y_{\ell,t} = \mathbf{F}_t' \boldsymbol{\theta}_{\ell,t} + u_{\ell,t}, \\ \boldsymbol{\theta}_{\ell,t} = \mathbf{H}_{\ell,t}\boldsymbol{\theta}_{\ell,t} + \mathbf{J}_{\ell,t}\, \boldsymbol{\theta}_{\ell\text{-}1,t} + \mathbf{w}_{\ell,t}, \end{cases} \tag{A.2}$$

where

$$\mathbf{H}_{\ell,t} = \begin{bmatrix} \mathbf{G}_t^{(11)} & \mathbf{G}_t^{(12)} \\ \mathbf{G}_t^{(21)} & \boldsymbol{\Gamma}_\ell \mathbf{G}_t^{(22)} \end{bmatrix}, \quad \mathbf{J}_{\ell,t} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{I}-\boldsymbol{\Gamma}_\ell) \end{bmatrix},$$

$$\mathbf{G}_t = \begin{bmatrix} \mathbf{G}_t^{(11)} & \mathbf{G}_t^{(12)} \\ \mathbf{G}_t^{(21)} & \mathbf{G}_t^{(22)} \end{bmatrix} \quad \text{partitioned corresponding to } \boldsymbol{\theta}_{\cdot,t} = \begin{bmatrix} \boldsymbol{\theta}_{\cdot,t}^{(1)} \\ \boldsymbol{\theta}_{\cdot,t}^{(2)} \end{bmatrix},$$

$\boldsymbol{\Gamma}_\ell$ is a diagonal matrix with entries taking values between 0 and 1, when ℓ=1, $\boldsymbol{\Gamma}_1 = \mathbf{I}$.

From the second equation of model (A.2) turns out

$$\begin{aligned} \boldsymbol{\theta}_{\ell,t} &= \mathbf{H}_{\ell,t}\cdots\mathbf{H}_{\ell,1}\, \boldsymbol{\theta}_{\ell,0} + \mathbf{H}_{\ell,t}\cdots\mathbf{H}_{\ell,2}\, \mathbf{J}_{\ell,1}\, \boldsymbol{\theta}_{\ell\text{-}1,1} + \cdots + \mathbf{J}_{\ell,t}\, \boldsymbol{\theta}_{\ell\text{-}1,t} \\ &+ \mathbf{H}_{\ell,t}\cdots\mathbf{H}_{\ell,2}\, \mathbf{w}_{\ell,1} + \mathbf{H}_{\ell,t}\cdots\mathbf{H}_{\ell,3}\, \mathbf{w}_{\ell,2} + \cdots + \mathbf{w}_{\ell,t} \\ &= [\mathbf{H}_{\ell,t}\cdots\mathbf{H}_{\ell,1}, \mathbf{H}_{\ell,t}\cdots\mathbf{H}_{\ell,2}\mathbf{J}_{\ell,1}, \cdots, \mathbf{J}_{\ell,t}, \mathbf{0}, \cdots \mathbf{0}]\, \boldsymbol{\Theta}_{\ell\text{-}1} + (\mathbf{H}_{\ell,t}\cdots\mathbf{H}_{\ell,1})\, \mathbf{s}_{\ell,0} \\ &+ [\mathbf{H}_{\ell,t}\cdots\mathbf{H}_{\ell,1}, \mathbf{H}_{\ell,t}\cdots \mathbf{H}_{\ell,2}, \cdots, \mathbf{H}_{\ell,t\text{-}1}\, \mathbf{I}, \mathbf{0}, \cdots, \mathbf{0}] \begin{bmatrix} \mathbf{0} \\ \mathbf{W}_{\ell,n} \end{bmatrix}. \end{aligned}$$

Therefore, $\boldsymbol{\Theta}_\ell = [\boldsymbol{\theta}'_{\ell,0}\, , \boldsymbol{\theta}'_{\ell,1}\, , \cdots , \boldsymbol{\theta}'_{\ell,1}, \cdots , \boldsymbol{\theta}'_{\ell,\mathrm{n}}]' = \mathbf{L}_\ell\, \boldsymbol{\Theta}_{\ell\text{-}1} + \mathbf{S} + \mathbf{M}_\ell \begin{bmatrix} \mathbf{0} \\ \mathbf{W}_{\ell,\mathrm{n}} \end{bmatrix}$, with $\mathbf{L}_\ell$, $\mathbf{S}$, $\mathbf{M}_\ell$ specified as

$$\mathbf{L}_\ell = \begin{bmatrix} \mathbf{I} & & & \\ \mathbf{H}_{\ell,1} & \mathbf{J}_{\ell,1} & & \\ \mathbf{H}_{\ell,2}\mathbf{H}_{\ell,1} & \mathbf{H}_{\ell,2}\mathbf{J}_{\ell,1} & \mathbf{J}_{\ell,2} & \\ \vdots & \vdots & \vdots & \ddots \\ \mathbf{H}_{\ell,\mathrm{n}}\cdots\mathbf{H}_{\ell,1} & \mathbf{H}_{\ell,\mathrm{n}}\cdots\mathbf{H}_{\ell,2}\mathbf{J}_{\ell,1} & \mathbf{H}_{\ell,\mathrm{n}}\cdots\mathbf{H}_{\ell,3}\mathbf{J}_{\ell,2}\cdots & \mathbf{J}_{\ell,\mathrm{n}} \end{bmatrix},$$

$$\mathbf{M}_\ell = \begin{bmatrix} \mathbf{I} & & & \\ \mathbf{H}_{\ell,1} & \mathbf{I} & & \\ \mathbf{H}_{\ell,2}\,\mathbf{H}_{\ell,1} & \mathbf{H}_{\ell,2} & \mathbf{I} & \\ \vdots & \vdots & \vdots & \ddots \\ \mathbf{H}_{\ell,\mathrm{n}}\cdots\mathbf{H}_{\ell,1} & \mathbf{H}_{\ell,\mathrm{n}}\cdots\mathbf{H}_{\ell,2} & \mathbf{H}_{\ell,\mathrm{n}}\cdots\mathbf{H}_{\ell,3} & \cdots\ \mathbf{I} \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} \mathbf{I} \\ \mathbf{H}_{\ell,1} \\ \mathbf{H}_{\ell,2}\mathbf{H}_{\ell,1} \\ \vdots \\ \mathbf{H}_{\ell,\mathrm{n}}\cdots\mathbf{H}_{\ell,1} \end{bmatrix} \mathbf{s}_{\ell,0};$$

Thus under scenario A, i.e.

$$(\boldsymbol{\Theta}_{\ell\text{-}1}\,|\,\bar{\underline{\mathrm{y}}}_{\ell\text{-}1}) \sim \mathrm{N}(\mathbf{m}_{\ell\text{-}1}, \mathbf{C}_{\ell\text{-}1}),\ \ \underline{\mathrm{u}}_{\ell,\mathrm{n}} \sim \mathrm{N}(\mathbf{0}, \mathbf{U}_{\ell,\mathrm{n}}),\ \ \underline{\mathbf{w}}_{\ell,\mathrm{n}} \sim \mathrm{N}(\mathbf{0}, \mathbf{W}_{\ell,\mathrm{n}}),$$

$$(\boldsymbol{\theta}_{\ell,0}\,|\,\bar{\underline{\mathrm{y}}}_{\ell\text{-}1}) = (\boldsymbol{\theta}_{\ell\text{-}1,0}\,|\,\bar{\underline{\mathrm{y}}}_{\ell\text{-}1}) + \mathbf{s}_{\ell,0},\ \ \mathbf{s}_{\ell,0} \text{ is a known constant vector,}$$

$$\{\mathrm{u}_{\ell,\mathrm{t}}\}, \{\mathbf{w}_{\ell,\mathrm{t}}\} \text{ and } \{\boldsymbol{\Theta}_{\ell\text{-}1}\,|\,\bar{\underline{\mathrm{y}}}_{\ell\text{-}1}\} \text{ are mutually independent,}$$

applying proposition 1, we obtain the prior joint distribution

$$\begin{bmatrix} \boldsymbol{\Theta}_\ell \\ \underline{\mathrm{Y}}_\ell^{(1)} \\ \underline{\mathrm{Y}}_\ell^{(2)} \end{bmatrix}\Bigg|\,\bar{\underline{\mathrm{y}}}_{\ell\text{-}1} \sim \mathrm{N}\left[\begin{bmatrix} \mathbf{m}_{\theta_\ell} \\ \mathbf{m}_{\mathrm{y}_{\ell.1}} \\ \mathbf{m}_{\mathrm{y}_{\ell,2}} \end{bmatrix}, \begin{bmatrix} \mathrm{C}_{\theta_\ell} & \mathrm{C}_{\theta_\ell \mathrm{y}_{\ell.1}} & \mathrm{C}_{\theta_\ell \mathrm{y}_{\ell.2}} \\ \mathrm{C}_{\mathrm{y}_{\ell.1}\theta_\ell} & \mathrm{C}_{\mathrm{y}_{\ell.1}} & \mathrm{C}_{\mathrm{y}_{\ell.12}} \\ \mathrm{C}_{\mathrm{y}_{\ell.2}\theta_\ell} & \mathrm{C}_{\mathrm{y}_{\ell.21}} & \mathrm{C}_{\mathrm{y}_{\ell.2}} \end{bmatrix}\right],$$

where $$\mathbf{m}_{\theta_\ell} = \mathbf{L}_\ell \ \mathbf{m}_{\ell\text{-}1} + \mathbf{s}_\ell, \quad \begin{pmatrix} \mathbf{m}_{y_{\ell.1}} \\ \mathbf{m}_{y_{\ell.2}} \end{pmatrix} = \mathbf{F}' \ \mathbf{m}_{\theta_\ell},$$

$$\mathbf{C}_{\theta_\ell} = \mathbf{L}_\ell \, \mathbf{C}_{\ell\text{-}1} \, \mathbf{L}'_\ell + \ \mathbf{M}_\ell \begin{bmatrix} \mathbf{0} & \\ & \mathbf{W}_{\ell,\mathrm{n}} \end{bmatrix} \mathbf{M}'_\ell \, ,$$

$$\begin{pmatrix} \mathbf{C}_{y_{\ell.1}\theta_\ell} \\ \mathbf{C}_{y_{\ell.2}\theta_\ell} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{\theta_\ell y_{\ell.1}} & \mathbf{C}_{\theta_\ell y_{\ell.2}} \end{pmatrix}' = \mathbf{F}'\mathbf{C}_{\theta_\ell} \, ,$$

$$\begin{pmatrix} \mathbf{C}_{y_{\ell.1}} & \mathbf{C}_{y_{\ell.12}} \\ \mathbf{C}_{y_{\ell.21}} & \mathbf{C}_{y_{\ell.2}} \end{pmatrix} = \mathbf{F}'\mathbf{C}_{\theta_\ell} \mathbf{F} + \mathbf{U}_{\ell,\mathrm{n}} \, .$$

Applying proposition 2 to above prior joint distribution yields the conditional distribution of $\boldsymbol{\Theta}_\ell$ and $\underset{\sim}{\mathrm{Y}}_\ell^{(2)}$ given observations $\underset{\sim}{\mathrm{y}}_{\ell,\mathrm{t}}$ as

$$\begin{pmatrix} \boldsymbol{\Theta}_\ell \\ \underset{\sim}{\mathrm{Y}}_\ell^{(2)} \end{pmatrix} | \, \bar{\underset{\sim}{\mathrm{y}}}_{\ell\text{-}1}, \underset{\sim}{\mathrm{y}}_{\ell,\mathrm{t}} \Bigg) \sim \mathrm{N}\left(\begin{pmatrix} \tilde{\mathbf{m}}_{\theta_\ell} \\ \tilde{\mathbf{m}}_{y_{\ell.2}} \end{pmatrix}, \begin{pmatrix} \tilde{\mathbf{C}}_{\theta_\ell} & \tilde{\mathbf{C}}_{\theta_\ell y_{\ell.2}} \\ \tilde{\mathbf{C}}_{y_{\ell.2}\theta_\ell} & \mathbf{C}_{y_{\ell.2}} \end{pmatrix} \right),$$

where $$\tilde{\mathbf{m}}_{\theta_\ell} = \mathbf{m}_{\theta_\ell} + \mathbf{C}_{\theta_\ell y_{\ell.1}} (\mathbf{C}_{y_{\ell.1}})^{-1} (\underset{\sim}{\mathrm{y}}_{\ell,\mathrm{t}} - \mathbf{m}_{y_{\ell.1}}),$$

$$\tilde{\mathbf{m}}_{y_{\ell.2}} = \mathbf{m}_{y_{\ell.2}} + \mathbf{C}_{y_{\ell.21}} (\mathbf{C}_{y_{\ell.1}})^{-1} (\underset{\sim}{\mathrm{y}}_{\ell,\mathrm{t}} - \mathbf{m}_{y_{\ell.1}}),$$

$$\begin{pmatrix} \tilde{\mathbf{C}}_{\theta_\ell} & \tilde{\mathbf{C}}_{\theta_\ell y_{\ell.2}} \\ \tilde{\mathbf{C}}_{y_{\ell.2}\theta_\ell} & \tilde{\mathbf{C}}_{y_{\ell.2}} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{\theta_\ell} & \mathbf{C}_{\theta_\ell y_{\ell.2}} \\ \mathbf{C}_{y_{\ell.2}\theta_\ell} & \mathbf{C}_{y_{\ell.2}} \end{pmatrix} - \begin{pmatrix} \mathbf{C}_{\theta_\ell y_{\ell.1}} \\ \mathbf{C}_{y_{\ell.21}} \end{pmatrix} \mathrm{C}_{y_{\ell.1}}^{-1} \begin{pmatrix} \mathbf{C}_{y_{\ell.1}\theta_\ell} & \mathbf{C}_{y_{\ell.12}} \end{pmatrix}.$$

In particular, we have the following joint predictive distribution:

$$(\underset{\sim}{Y}_\ell^{(2)} \mid \bar{\underset{\sim}{y}}_{\ell\text{-}1}, \underset{\sim}{y}_\ell^{(1)}) \sim N(\tilde{m}_{y_{\ell.2}}, \tilde{C}_{y_{\ell.2}}).$$

Under scenario B, i.e. $(\boldsymbol{\Theta}_{\ell\text{-}1} \mid \bar{\underset{\sim}{y}}_{\ell\text{-}1}, U) \sim N(\mathbf{m}_{\ell\text{-}1}, U\mathbf{C}_{\ell\text{-}1})$,

$$(\underset{\sim}{u}_{\ell,n} \mid U, \bar{\underset{\sim}{y}}_{\ell\text{-}1}) \sim N(\mathbf{0},\ U\mathbf{U}_{\ell,n}),\quad (\underset{\sim}{\mathbf{w}}_{\ell,n} \mid \bar{\underset{\sim}{y}}_{\ell\text{-}1}, U) \sim N(\mathbf{0},\ U\mathbf{W}_{\ell,n}),$$

$$(U \mid \bar{\underset{\sim}{y}}_{\ell\text{-}1}) \sim \mathbf{IG}\left(\frac{n_{\ell,0}}{2},\ \frac{d_{\ell,0}}{2}\right),$$

given U sequences $\{u_{\ell,t}\}$, $\{w_{\ell,t}\}$ and $\{\boldsymbol{\Theta}_{\ell\text{-}1,t} \mid \bar{\underset{\sim}{y}}_{\ell\text{-}1}\}$ are mutually independent,

we have the following

$$\left[\begin{array}{c|c} \boldsymbol{\Theta}_\ell & \\ \underset{\sim}{Y}_\ell^{(1)} & \bar{\underset{\sim}{y}}_{\ell\text{-}1}, U \\ \underset{\sim}{Y}_\ell^{(2)} & \end{array}\right] \sim N\left[\begin{bmatrix} m_{\theta_\ell} \\ m_{y_{\ell.1}} \\ m_{y_{\ell.2}} \end{bmatrix}, U\begin{bmatrix} C_{\theta_\ell} & C_{\theta_\ell y_{\ell.1}} & C_{\theta_\ell y_{\ell.2}} \\ C_{y_{\ell.1}\theta_\ell} & C_{y_{\ell.1}} & C_{y_{\ell.12}} \\ C_{y_{\ell.2}\theta_\ell} & C_{y_{\ell.21}} & C_{y_{\ell.2}} \end{bmatrix}\right],$$

$$\left[\begin{array}{c|c} \boldsymbol{\Theta}_\ell & \\ \underset{\sim}{Y}_\ell^{(1)} & \bar{\underset{\sim}{y}}_{\ell\text{-}1} \\ \underset{\sim}{Y}_\ell^{(2)} & \end{array}\right] \sim T_{n_{\ell,0}}\left[\begin{bmatrix} m_{\theta_\ell} \\ m_{y_{\ell.1}} \\ m_{y_{\ell.2}} \end{bmatrix}, \frac{d_{\ell,0}}{n_{\ell,0}}\begin{bmatrix} C_{\theta_\ell} & C_{\theta_\ell y_{\ell.1}} & C_{\theta_\ell y_{\ell.2}} \\ C_{y_{\ell.1}\theta_\ell} & C_{y_{\ell.1}} & C_{y_{\ell.12}} \\ C_{y_{\ell.2}\theta_\ell} & C_{y_{\ell.21}} & C_{y_{\ell.2}} \end{bmatrix}\right],$$

$$(U \mid \bar{y}_{\ell-1}, y_{\ell,t}) \sim IG\left(\frac{n_{\ell,t}}{2}, \frac{d_{\ell,t}}{2}\right),$$

$$\left(\begin{matrix}\Theta_\ell \\ Y_\ell^{(2)}\end{matrix} \mid \bar{y}_{\ell-1}, y_{\ell,t}, U\right) \sim N\left(\left(\begin{matrix}\tilde{m}_{\theta_\ell} \\ \tilde{m}_{y_{\ell.2}}\end{matrix}\right), \; U \; \left(\begin{matrix}\tilde{C}_{\theta_\ell} & \tilde{C}_{\theta_\ell y_{\ell.2}} \\ \tilde{C}_{y_{\ell.2}\theta_\ell} & \tilde{C}_{y_{\ell.2}}\end{matrix}\right)\right),$$

$$\left(\begin{matrix}\Theta_\ell \\ Y_\ell^{(2)}\end{matrix} \mid \bar{y}_{\ell-1}, \; y_{\ell,t}\right) \sim T_{n_{\ell,t}}\left(\left(\begin{matrix}\tilde{m}_{\theta_\ell} \\ \tilde{m}_{y_{\ell.2}}\end{matrix}\right), \; \frac{d_{\ell,t}}{n_{\ell,t}} \left(\begin{matrix}\tilde{C}_{\theta_\ell} & \tilde{C}_{\theta_\ell y_{\ell.2}} \\ \tilde{C}_{y_{\ell.2}\theta_\ell} & \tilde{C}_{y_{\ell.2}}\end{matrix}\right)\right),$$

where $\quad n_{\ell,t} = n_{\ell,0} + t, \qquad d_{\ell,t} = d_{\ell,0} + d'_{\ell,t},$

$$d'_{\ell,t} = (y_{\ell,t} - m_{y_{\ell.1}})' (C_{y_{\ell.1}})^{-1} (y_{\ell,t} - m_{y_{\ell.1}}),$$

all submatrices are specified as the same as under scenario A.

REFERENCES

Habibi, A. (1972). "Two-dimensional Bayesian estimate of images," Proc. IEEE, V60, 878-883.

Harrison, P. J. and Stevens, C. F. (1971). "A Bayesian approach to short-term forecasting," Oper. Res. Quarterly, V22, 341-362.

Harrison, P. J. and Stevens, C. F. (1976). "Bayesian Forecasting," JRSS, B. V38, 205-247.

Kalman, R. E. (1960). "A new approach to linear filtering and prediction problems," J. of Basic Engineering, V82, 35-45.

Katayama, T. and Kosaka, M. (1979). "Recursive filtering algorithm for a two-dimensional system." IEEE Trans. Automat. Contr. V. AC-24, 130-132.

Smith, A. F. M. and West, M. (1983). "Monitoring rental transplants: an application of the multi-process Kalman filter," Biometrics, V.39, 867-878.

West, M. and Harrison, J. (1989). Bayesian Forecasting and Dynamic Models. Spring-Verlag.

Woods, J. W. and Ingle, V. K. (1981). "Kalman filtering in two dimensions: further results." IEEE Trans. Acoustic, Speech, and Signal Processing, V. ASSP-29, 188-197.

Woods, J. W. and Radewan, C. H. (1977). "Kalman filtering in two dimensions," IEEE Trans. Inform. Theory, V.IT-23, 473-482.